

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE INFORMÁTICA

DEPARTAMENTO DE INGENIERÍA DEL SOFTWARE E INTELIGENCIA ARTIFICIAL



TESIS DOCTORAL

**SISTEMA DE VISIÓN PARA AGRICULTURA DE PRECISIÓN:
IDENTIFICACIÓN EN TIEMPO REAL DE LÍNEAS DE CULTIVO Y MALAS HIERBAS
EN CAMPOS DE MAÍZ**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

José Miguel Guerrero Hernández

Directores

Gonzalo Pajares Martinsanz
María Guijarro Mata-García
Ángela Ribeiro Seijas

Madrid, 2015

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE INFORMÁTICA

Departamento de Ingeniería del Software e Inteligencia Artificial



SISTEMA DE VISIÓN PARA AGRICULTURA DE PRECISIÓN: IDENTIFICACIÓN EN TIEMPO REAL DE LÍNEAS DE CULTIVO Y MALAS HIERBAS EN CAMPOS DE MAÍZ

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

José Miguel Guerrero Hernández

Bajo la dirección de los doctores

Gonzalo Pajares Martinsanz

María Guijarro Mata-García

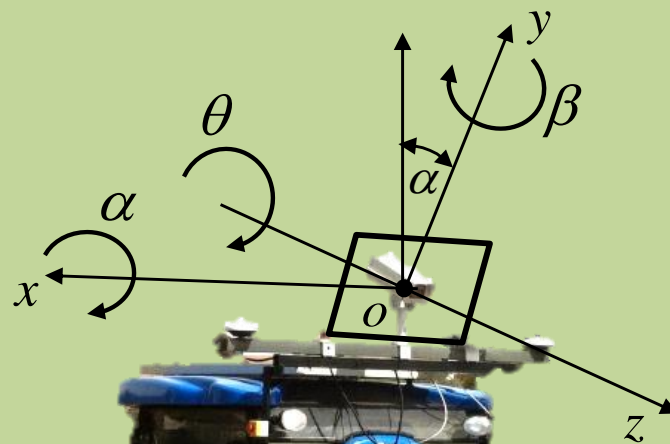
Ángela Ribeiro Seijas

Madrid, 2014

SISTEMA DE VISIÓN PARA AGRICULTURA DE PRECISIÓN: IDENTIFICACIÓN EN TIEMPO REAL DE LÍNEAS DE CULTIVO Y MALAS HIERBAS EN CAMPOS DE MAÍZ

TESIS DOCTORAL

José Miguel Guerrero Hernández



Universidad Complutense de Madrid
Facultad de Informática
Departamento de Ingeniería del Software e Inteligencia Artificial

SISTEMA DE VISIÓN PARA AGRICULTURA DE PRECISIÓN: IDENTIFICACIÓN EN TIEMPO REAL DE LÍNEAS DE CULTIVO Y MALAS HIERBAS EN CAMPOS DE MAÍZ

TESIS DOCTORAL



José Miguel Guerrero Hernández

Departamento de Ingeniería del Software e Inteligencia Artificial
Facultad de Informática
Universidad Complutense de Madrid

Madrid 2014

Departamento de Ingeniería del Software e Inteligencia Artificial
Facultad de Informática
Universidad Complutense de Madrid

Tesis Doctoral

Sistema de visión para agricultura de precisión: identificación en tiempo real de líneas de cultivo y malas hierbas en campos de maíz

Autor

José Miguel Guerrero Hernández

Directores

Dr. Gonzalo Pajares Martinsanz

Dra. María Guijarro Mata-García

Dra. Ángela Ribeiro Seijas

RESUMEN

El control de malas hierbas en grandes extensiones de terreno resulta costoso y a veces contaminante. El avance en los últimos años de los sistemas de control y automatización en agricultura ha favorecido la aparición del concepto de Agricultura de Precisión (AP), para intervenir sobre el cultivo, en el momento adecuado y en el lugar preciso. La presente investigación tiene su origen principalmente en dos proyectos, RHEA y AGL2011, en los cuales la identificación de líneas de cultivo en maíz y la discriminación entre cultivo y mala hierba suscita un interés económico, ecológico y comercial de cara a su implantación en tractores autónomos.

El principal aporte presentado en esta memoria de tesis consiste en el diseño de un sistema experto de visión por computador capaz de identificar en tiempo real las líneas de cultivo y las malas hierbas en campos de maíz. Además, el sistema propuesto controla el guiado del tractor y el solapamiento de las zonas de tratamiento con el fin de aplicar un tratamiento selectivo.

Como estrategia a seguir se propone una arquitectura abierta del sistema dividida en dos módulos. El primero, se encarga de capturar una imagen de calidad y de segmentar dicha imagen de manera que se identifiquen las plantas tanto de cultivo como de mala hierba. Para ello, se aplican: *a)* técnicas para controlar de forma automática el tiempo de exposición de la cámara para poder hacer frente a las condiciones cambiantes de iluminación de los entornos de exterior; *b)* técnicas para corregir las imágenes en caso de verse alteradas por filtros de corte de las radiaciones ultravioleta e infrarroja; *c)* métodos de segmentación y umbralización específicamente diseñados para la identificación de las plantas. El segundo módulo, recibe la imagen segmentada procedente del módulo anterior e identifica las líneas de cultivo y malas hierbas mediante: *a)* estrategias robustas para la identificación precisa de líneas de cultivo y malas hierbas aplicando la geometría conocida del sistema de visión a partir de sus parámetros intrínsecos y extrínsecos; *b)* técnicas para el control del solapamiento y el guiado preciso del tractor.

Finalmente se presenta el sistema de visión operativo dentro de uno de los tractores prototipo que conforman la flota de vehículos del proyecto RHEA, cuya demostración final tuvo lugar el 21 de mayo de 2014 con resultados satisfactorios.

Palabras clave

Agricultura de precisión, Automatización, Detección de líneas, Detección de malas hierbas, Guiado, Maíz, Procesamiento de imágenes, Segmentación, Sistemas expertos, Solapamiento, Tiempo de exposición, Tiempo real, Tractor autónomo, Tratamientos selectivos por zonas, Umbralización, Visión artificial.

ABSTRACT

The control of weeds in large land areas results expensive and most times pollutant. In the last years, important advances through the application of control systems and automation in agriculture have been made possible the appearance of Precision Agriculture (PA), which means to carry out the right intervention, in the right place at the right time. The research developed resides in two projects, RHEA and AGL2011, in which the crop rows identification in maize fields and discrimination between crop and weeds cause a financial, ecological and commercial interest due to the possible installation in autonomous tractors.

The main contribution presented in this thesis is the design of an expert system based on computer vision for real-time crop rows and weeds identification in maize fields. Furthermore, the proposed system controls the guidance of the tractor and the overlapping of the areas of treatment in order to apply a site-specific treatment.

An open architecture system is proposed, divided into two modules. The first is responsible for capturing a quality image and for obtaining a segmented image where the crop and weeds plants are identified. To do this, we apply: *a)* automatic techniques to adjust the exposure time of the camera, to cope with different light changes caused by work in outdoor environments; *b)* techniques for image correction, which are modified by an ultra-violet and infrared cut filter; *c)* segmentation and thresholding methods specifically designed for plant identification. The second module receives the segmented image from the previous module and identifies the crop rows and weeds by: *a)* robust strategies for the precise identification of crop rows and weeds applying the known geometry of the vision system from its intrinsic and extrinsic parameters, previously calibrated; *b)* techniques to control overlapping and the precise guidance of the tractor.

Finally, the implementation of the proposed vision system is provided an also the integration within one of the prototype tractors belonging to the fleet of agricultural vehicles in the RHEA project, which final demo took place in May 21, 2014 with successful results.

Keywords

Automation, Autonomous tractor, Computer vision, Expert systems, Exposure time, Guidance, Image processing, Lines detection, Maize, Overlapping, Precision agriculture, Real-time, Segmentation, Site-specific treatments, Thresholding, Weeds detection.

Nunca dejes que nadie te diga que no puedes hacer algo. Ni siquiera yo. Si tienes un sueño, tienes que protegerlo. Las personas que no son capaces de hacer algo te dirán que tú tampoco puedes. Si quieres algo ve por ello, y punto.

[En busca de la felicidad]

A mis padres, María Rosa y José Miguel

A Verónica

Agradecimientos

*Nunca vayas por el camino trazado,
porque conduce hacia donde otros han ido ya.*

Alexander Graham Bell

Es imposible no echar la vista atrás y ver lo rápido que han pasado los años. Ha sido un largo y arduo camino, pero al fin llega el fruto de toda la dedicación y trabajo que tantas horas de sueño se ha llevado.

Es por ello que como no podía ser de otra manera, quisiera agradecer en primer lugar a mi director, Gonzalo Pajares, todo el trabajo, tiempo y dedicación que me ha ofrecido. Gracias por haberme dado la oportunidad de formar parte de este proyecto, y sobre todo, gracias por la confianza que depositaste en mí.

Agradecer a mi codirectora, María Guijarro, el haber despertado mi curiosidad por la investigación cuando aún estaba acabando la carrera. Gracias por tu comprensión, apoyo y consejos, que sin duda me han sido y serán de mucha ayuda.

También, gracias a Ángela Ribeiro por su codirección y consejos, y a todos los miembros del Centro de Automática y Robótica del Consejo Superior de Investigaciones Científicas que tantas horas hemos pasado juntos, en especial a Pablo Gonzalez de Santos, Luis y Mariano, porque al final, el trabajo y esfuerzo de todos ha hecho posible que esto saliera adelante.

Gracias a Jesús Manuel de la Cruz y al resto de miembros del grupo multidisciplinar ISCAR de la Universidad Complutense, por haberme permitido formar parte él.

En un plano más formal, agradecer a la Universidad Complutense de Madrid por la beca predoctoral concedida que me permitió iniciar esta investigación, así como al Ministerio de Educación, Cultura y Deporte del gobierno de España por el contrato del programa de Formación del Profesorado Universitario (FPU) que me ha permitido finalizarla.

Durante todo este tiempo no sólo he crecido profesionalmente, sino también como persona, es por eso que sin dudarlo quiero agradecer a toda esa gente que ha hecho posible que desconectara del trabajo y con la que tan buenos ratos he pasado durante estos años: Isa, Juanki, Rosa, Santi, José, David y Antonio, mil gracias. Y por supuesto a Martín, porque han sido varios años trabajando juntos y él sabe que todo esto no habría sido posible sin su ayuda.

Y por último, quiero dar mi más profundo agradecimiento a mi familia, por su amor incondicional desde el inicio; a mis padres, por su esfuerzo, ternura y sacrificio; y a Verónica, por haber estado a mi lado y no haber dejado que me rindiera nunca. En definitiva, gracias por haberme aguantado durante todos estos años y haber creído en mí.

¡Muchas gracias!

Índice

CAPÍTULO 1. INTRODUCCIÓN	1
1.1 Antecedentes.....	1
1.2 Identificación del problema.....	6
1.3 Motivación y objetivos	7
1.3.1 Motivación	7
1.3.2 Objetivos	10
1.4 Metodología.....	11
1.5 Aportes a la investigación	13
1.5.1 Publicaciones en Revistas Indexadas	13
1.5.2 Publicaciones en Revistas no Indexadas	16
1.5.3 Congresos y conferencias.....	17
1.5.4 Seminarios.....	20
1.5.5 Emisiones radiofónicas	21
1.6 Organización de la tesis.....	21
 CAPÍTULO 2. ESTADO DEL ARTE	 23
2.1 Visión general.....	23
2.2 Cámaras	24
2.2.1 Sensor CCD.....	24
2.2.2 Sensor CMOS.....	25
2.2.3 Comparativa de sensores CCD y CMOS	26
2.2.4 Formación de la imagen a color	28
2.3 Sistema óptico	29
2.3.1 Distancia focal de la lente	30
2.3.2 Iris: fijo, manual y automático	31
2.3.3 Enfoque	33

2.4 Técnicas de segmentación	33
2.4.1 Métodos de segmentación	34
2.4.2 Revisión de métodos de segmentación en imágenes naturales	36
2.4.3 Umbralización de imágenes	39
2.5 Técnicas de detección de líneas de cultivo y malas hierbas	40
2.5.1 Métodos basados en la exploración de líneas horizontales	41
2.5.2 Métodos basados en la transformada de Hough	42
2.5.3 Métodos basados en el punto de fuga.....	43
2.5.4 Métodos basados en visión estéreo	43
2.5.5 Métodos basados en regiones	44
2.5.6 Métodos basados en la acumulación de plantas verdes.....	44
2.5.7 Métodos basados en análisis de frecuencia	45
2.5.8 Métodos basados en regresión lineal.....	46

CAPÍTULO 3. PROCESAMIENTO DE IMÁGENES EN AGRICULTURA DE PRECISIÓN.....47

3.1 Introducción.....	47
3.2 Imágenes: modelos de color y efecto <i>vignetting</i>	49
3.2.1 Modelos de color.....	49
3.2.1.1 RGB.....	49
3.2.1.2 HSI	50
3.2.1.3 YUV	51
3.2.1.4 $YD_B D_R$	51
3.2.2 Efecto <i>vignetting</i>	52
3.3 Extracción del verde	55
3.4 Umbralización	59
3.4.1 Umbralización manual	60
3.4.2 Método de Otsu	61
3.4.3 Operaciones morfológicas.....	63
3.4.3.1 Erosión	64
3.4.3.2 Dilatación	65
3.4.3.3 Apertura y cierre.....	66
3.5 Identificación de líneas de cultivo	67
3.5.1 Regresión lineal simple	68
3.5.2 Estimador Theil-Sen.....	70

CAPÍTULO 4. SEGMENTACIÓN DE IMÁGENES AGRÍCOLAS EN CAMPOS DE MAÍZ73

4.1 Introducción.....	73
4.2 Identificación y tratamiento de imágenes con alta intensidad luminosa	73
4.2.1 Identificación de imágenes con alta intensidad luminosa	74
4.2.1.1 Fase de aprendizaje: ajustes del umbral	76
4.2.1.2 Fase de decisión y procesamiento	77
4.2.2 Tratamiento de imágenes	80
4.2.2.1 Cámaras con iris manual y que no permiten un ajuste del tiempo de exposición.....	81
4.2.2.2 Cámaras que permiten un ajuste del tiempo de exposición.....	86
4.2.3 Proceso completo para el tratamiento de imágenes.....	93
4.3 Identificación de plantas y binarización basada en $YD_B D_R$	96
4.4 Segmentación e identificación de plantas enmascaradas y sin enmascarar mediante Máquinas de Vectores Soporte	101
4.4.1 Análisis de índices vegetativos.....	101
4.4.2 Planteamiento del enfoque	102
4.4.3 Fase de aprendizaje basada en Máquinas de Vectores Soporte.....	104
4.4.4 Fase de decisión	109

CAPÍTULO 5. SISTEMA EXPERTO: DETECCIÓN DE LÍNEAS DE CULTIVO Y MALAS HIERBAS, GUIADO SOLAPAMIENTO113

5.1 Introducción.....	113
5.2 Arquitectura del sistema experto automático	116
5.3 Segmentación de la imagen	118
5.4 Estimación de las líneas de cultivo.....	120
5.4.1 Trazado de las líneas de cultivo estimadas mediante geometría	121
5.4.2 Posicionamiento del sistema: detección de la orientación	123
5.4.3 Corrección de las líneas estimadas: estimador Theil-Sen	128
5.5 Detección de densidades de malas hierbas	132
5.6 Guiado del tractor	135
5.7 Solapamiento de imágenes	140

CAPÍTULO 6. ANÁLISIS DE RESULTADOS143

6.1 Introducción.....	143
6.2 Características de los componentes del sistema de visión artificial	144
6.2.1 Cámaras.....	144
6.2.2 Lentes	147
6.2.3 Unidad de medición inercial.....	149
6.2.4 Plataformas de procesamiento.....	150
6.3 Análisis de la segmentación de imágenes	151
6.3.1 Identificación de imágenes con alta intensidad lumínica	151
6.3.2 Vignetting y balance de blancos.....	153
6.3.3 Ajuste del tiempo de exposición	155
6.3.3.1 Precisión en la detección de verdes	157
6.3.3.2 Tendencia a adquirir imágenes subexpuestas, normales o sobreexpuestas ...	159
6.3.4 Identificación de plantas y binarización basada en $YD_B D_R$	160
6.3.5 Segmentación e identificación de plantas enmascaradas y sin enmascarar mediante SVM.....	161
6.4 Análisis del sistema experto	166
6.4.1 Estimación de líneas de cultivo: detección de la orientación	166
6.4.2 Corrección de líneas mediante el estimador Theil-Sen	169
6.4.2.1 Análisis cualitativo	171
6.4.2.2 Análisis cuantitativo	173
6.4.3 Guiado del tractor.....	177
6.4.4 Solapamiento de imágenes	178
6.5 Sistema experto desarrollado para el proyecto RHEA	179

CAPÍTULO 7. CONCLUSIONES Y TRABAJO FUTURO185

7.1 Introducción.....	185
7.2 Conclusiones.....	185
7.2.1 Segmentación de imágenes	186
7.2.2 Detección de líneas de cultivo.....	187
7.2.3 Guiado	188
7.2.4 Solapamiento.....	188
7.3 Trabajo futuro	189

CAPÍTULO 8. ENGLISH SUMMARY	193
8.1 Background.....	193
8.2 Problem statement	197
8.3 Motivation and goals	198
8.3.1 Motivation	198
8.3.2 Goals	200
8.4 Revision of methods	201
8.5 Support Vector Machines for masked and unmasked plants identification ...	202
8.5.1 Learning phase based on Support Vector Machines	203
8.5.2 Decision phase.....	206
8.6 Expert system: crop rows and weeds detection	206
8.6.1 Expert system architecture	208
8.6.2 Image segmentation.....	209
8.6.3 Crop rows estimation	210
8.6.3.1 Tracing expected crop rows	210
8.6.3.2 Correction of the expected crop rows: Theil-Sen estimator	211
8.6.4 Weeds density detection.....	213
8.7 Results	213
8.7.1 Support Vector Machines for masked and unmasked plants identification	214
8.7.2 Crop rows estimation: Theil-Sen estimator.....	217
8.7.2.1 Qualitative analysis	219
8.7.2.2 Quantitative analysis	220
8.8 Conclusions	223
8.8.1 Image segmentation.....	223
8.8.2 Crop rows detection	224
8.9 Future work	224
 BIBLIOGRAFÍA / BIBLIOGRAPHY	 227

Lista de Figuras

Figura 1.1	(a) UAV volando sobre un campo de maíz para detectar rodales de malas hierbas con el fin de generar un mapa de tratamientos específico por zonas. (b) Mapa de malas hierbas obtenido. (Imágenes cortesía de F. López-Granados y J.M. Peña; Instituto de Agricultura Sostenible, CSIC-Córdoba, España; Peña y col., 2012, 2013).	4
Figura 1.2	Sistema de visión artificial directo.	7
Figura 2.1	Esquema de un sensor CCD.	25
Figura 2.2	Esquema de un sensor CMOS.	26
Figura 2.3	Ejemplos de sensores con sus respectivos factores de multiplicación de la distancia focal, CDD a la izquierda y CMOS a la derecha.	27
Figura 2.4	Esquema del mosaico Bayer.	29
Figura 2.5	Relación entre distancias focales y ángulos de visión.	30
Figura 2.6	Ejemplo de sensor de 2/3" y lente de 1/2".	31
Figura 2.7	Ejemplos de sistemas ópticos, a la izquierda sistema con iris manual y a la derecha sistema con iris automático.	33
Figura 2.8	Cámara con sensor CCD, sistema óptico con iris manual y filtro de corte ultravioleta e infrarrojo.	33
Figura 3.1	Esquema del proceso general para la identificación de líneas de cultivo y malas hierbas.	48
Figura 3.2	Esquema del modelo de color RGB.	50

Figura 3.3	Esquema del modelo de color HSI.	51
Figura 3.4	(a) Incidencia de los rayos sin interposición de filtros o parasoles sobre la lente. (b) <i>Vignetting</i> mecánico, incidencia de los rayos sobre una lente con un parasol mal colocado. (c) <i>Vignetting</i> óptico, incidencia de los rayos sobre una lente con un filtro multicapa.	52
Figura 3.5	Imágenes con efecto <i>vignetting</i> : (a) mecánico, provocado por el uso de un filtro cuyo tamaño no se ajusta al sistema; (b) óptico, al aplicar un algoritmo de extracción de la capa vegetal.	53
Figura 3.6	(a) Imagen de líneas de cultivo tomada con la cámara SVS-VISTEK sin filtro de corte UV/IR; (b) la misma imagen con filtro de bloqueo UV/IR.	54
Figura 3.7	(a) Imagen original. (b) Extracción de verdes mediante el índice COM (Guijarro y col., 2011).	59
Figura 3.8	(a) Imagen en escala de grises obtenida a partir de la figura 3.7(a). (b) Histograma obtenido.	60
Figura 3.9	Binarización mediante el método de Otsu de la figura 3.7(b).	62
Figura 3.10	Erosión de la figura 3.9 utilizando el elemento estructural M_8 .	65
Figura 3.11	Dilatación de la figura 3.9 utilizando el elemento estructural M_8 .	66
Figura 3.12	(a) Apertura. (b) Cierre.	67
Figura 3.13	Ejemplo de recta obtenida por regresión lineal simple a partir de un conjunto de píxeles.	69
Figura 3.14	Ejemplo de recta obtenida por regresión lineal simple a partir de un conjunto de píxeles donde un grupo de píxeles de mala hierba forman parte del estudio.	70
Figura 3.15	Comparativa de una recta obtenida por regresión lineal simple (línea negra) y otra recta obtenida utilizando el estimador Theil-Sen (línea verde).	72
Figura 4.1	Esquema del proceso de aprendizaje. Conjuntos C1 y C2 con los datos característicos de cada uno, porcentaje de verde (PG), valor medio de intensidad (I) y umbral seleccionado (T).	77
Figura 4.2	Esquema del proceso de decisión para determinar la validez de la imagen capturada y pasos en el proceso de binarización.	79

Figura 4.3	(a) Imagen original con un nivel alto de iluminación. (b) Imagen original con un nivel adecuado de iluminación.	80
Figura 4.4	Imágenes obtenidas después de aplicar ExG y el método de Otsu. (a) A partir de la figura 4.3(a) con un alto nivel de intensidad, con valor medio de intensidad 0.876. (b) A partir de la figura 4.3(b) con un nivel adecuado de intensidad, con valor medio de intensidad 0.4688.	80
Figura 4.5	Filtro utilizado en la corrección del <i>vignetting</i> .	82
Figura 4.6	(a) Imagen donde se aprecia el <i>vignetting</i> en los laterales de la imagen y en las esquinas; (b) Imagen donde se aprecia la corrección del efecto <i>vignetting</i> al aplicar la corrección sobre el canal rojo.	84
Figura 4.7	(a) Imagen original; (b) Imagen modificada con el balance de blancos.	86
Figura 4.8	Esquema del proceso de ajuste del tiempo de exposición.	87
Figura 4.9	Imágenes consecutivas durante el movimiento del tractor autónomo. (a) Imagen con un tiempo de exposición bien ajustado. (b) Imagen con un tiempo de exposición excesivo.	90
Figura 4.10	Imagen donde puede verse la zona de interés (ROI).	90
Figura 4.11	(a) Imagen donde puede verse la zona de interés (ROI) delante del tractor. (b) Imagen tomada de la misma escena a través de la cámara montada en el tractor.	91
Figura 4.12	Esquema completo para la identificación de imágenes alta o insuficientemente iluminadas, su corrección, tratamiento y binarización para el posterior procesado.	95
Figura 4.13	Descomposición de una imagen RGB en $YD_B D_R$ (a) Imagen original; (b) Imagen de la luminancia (Y); (c) Imagen de crominancia azul (D_B); (d) Imagen de crominancia roja (D_R).	97
Figura 4.14	Imagen resultante al aplicar el nuevo método basado en $YD_B D_R$.	99
Figura 4.15	Imagen resultante al aplicar los filtros al nuevo método basado en $YD_B D_R$.	100

Figura 4.16	(a) Las malas hierbas de la zona central inferior y de la zona derecha aparecen como enmascaradas. (b) Las malas hierbas de la zona central aparecen enmascaradas. (c) El herbicida fue aplicado dos días antes (las malas hierbas en la zona central de las líneas de cultivo avanzan hacia una etapa de secado), además, el campo también recibió una lluvia directa.	103
Figura 4.17	Esquema del proceso de aprendizaje.	105
Figura 4.18	Distribución de patrones y vectores soporte en el espacio tridimensional RGB.	106
Figura 4.19	(a) Índice de vegetación con la estrategia combinada (COM); (b) Imagen binaria mediante Otsu que contiene las plantas enmascaradas y sin enmascarar.	110
Figura 4.20	Plantas sin enmascarar (marcas verdes) y enmascaradas (marcas rojas). (a) Imagen resultado a partir de la figura 4.16(b). (b) Imagen resultado a partir de la figura 4.16(c).	111
Figura 5.1	Sistema de referencia: parámetros extrínsecos.	114
Figura 5.2	Arquitectura del sistema experto automático.	118
Figura 5.3	(a) Imagen original. (b) Índice GA extraído de la imagen (a). (c) Imagen binaria tras aplicar la umbralización de Otsu.	120
Figura 5.4	Imagen original, figura 5.3(a), con la proyección de cuatro líneas de cultivo.	123
Figura 5.5	(a) Imagen original siguiendo una trayectoria desviada. (b) Proyección de líneas sobre la figura 5.5(a) calculadas en función de un ángulo de guiñada fijo.	126
Figura 5.6	Rango de búsqueda.	126
Figura 5.7	Proyección de líneas sobre la figura 5.5(a) con un ángulo de guiñada variable.	128
Figura 5.8	Correcciones de las líneas estimadas (líneas rojas) mediante el estimador Theil-Sen (líneas verdes). (a) Figura 5.4, siguiendo una trayectoria correcta. (b) Figura 5.7, siguiendo una trayectoria desviada.	131
Figura 5.9	Intersecciones (puntos rojos) de las bisectrices con las líneas horizontales a 25 cm de distancia. (a) Figura 5.4, siguiendo una trayectoria correcta. (b) Figura 5.7, siguiendo una trayectoria desviada.	133

Figura 5.10	Intersecciones (puntos rojos) de las bisectrices con las líneas horizontales a 25 cm de distancia y ancho (puntos amarillos) entre los cuales se considera que los verdes pertenecen al cultivo. (a) Figura 5.4, siguiendo una trayectoria correcta. (b) Figura 5.7, siguiendo una trayectoria desviada.	134
Figura 5.11	Ejemplo de la ROI con una división en cuadrículas de 25 cm de longitud.	135
Figura 5.12	Ejemplo de planificación sobre un campo.	136
Figura 5.13	Desplazamiento calculado de la figura 5.10(a).	138
Figura 5.14	Desplazamiento calculado de la figura 5.10(b).	139
Figura 6.1	HP Photosmart R817. (a) Vista frontal. (b) Vista trasera.	145
Figura 6.2	BASLER scA1400-17FC.	146
Figura 6.3	SVS-VISTEK SVS4050CFLGEA.	147
Figura 6.4	Lente PENTAX C30811KP.	147
Figura 6.5	Schneider Kreuznach Cinegon 1.9/10-090. (a) Lente sin adaptadores. (b) Lente con adaptadores para el filtro y la cámara.	148
Figura 6.6	(a) Filtro Schneider UV/IR 486. (b) Curva de respuesta del filtro Schneider UV/IR 486.	149
Figura 6.7	MicroStrain 3DM-GX3-35(AHRS) con GPS.	150
Figura 6.8	Sistema NI cRIO-9082 RT.	151
Figura 6.9	Variación del tiempo de exposición.	157
Figura 6.10	Variaciones del tiempo de exposición para los métodos RGB y HSI.	160
Figura 6.11	Ejemplo de corrección de líneas de cultivo mediante el estimador Theil-Sen (línea roja) y la regresión lineal simple (línea azul) considerando la dispersión de plantas (píxeles verdes).	170
Figura 6.12	(a) Líneas de cultivo idealmente estimadas calculadas mediante los parámetros intrínsecos y extrínsecos (líneas amarillas) y un margen de tolerancia del 50%. (b) Corrección de las líneas usando el estimador Theil-Sen (líneas rojas) y la regresión lineal simple (líneas azules). Las circunferencias verdes marcan las zonas con píxeles de malas hierbas.	171

Figura 6.13	Promedio de los tiempos de procesamiento para el proceso completo: extracción de verdes (GA), binarización (Otsu) y ajuste de líneas (Theil-Sen) para los tres tamaños de reducción y los tres márgenes de tolerancia.	176
Figura 6.14	Carcasa. (a) Vista exterior. (b) Vista interior.	180
Figura 6.15	Esquema del sistema de visión instalado en el prototipo del tractor RHEA.	180
Figura 6.16	Esquema de conexión de la plataforma de procesamiento cRIO-9082 con la IMU y la cámara que forman parte del sistema de visión.	181
Figura 6.17	Interfaz de configuración.	182
Figura 6.18	Interfaz de usuario. Vista de la aplicación durante el tratamiento.	183
Figura 6.19	Archivo de configuración.	184
Figura 7.1	Detección de líneas en carreteras de doble sentido sin líneas pintadas, en verde las detectadas, en rojo las estimadas y en amarillo los carriles identificados (Imágenes cortesía de AutoNOMOS labs, Universidad Libre de Berlín, Alemania). (a) Carretera simple de doble sentido. (b) Carretera con estacionamiento a la derecha.	190
Figura 7.2	Aplicación del sistema de visión propuesto para la detección de líneas de carretera (continuas y discontinuas) y bordillos o límites de la carretera (Imágenes cortesía de AutoNOMOS labs, Universidad Libre de Berlín, Alemania). (a) Imagen original. (b) Extracción de bordes. (c) Detección de líneas y su corrección (líneas verdes pertenecientes a los bordes del carril, líneas o bordillos) e identificación de carriles (líneas amarillas).	191

Lista de Tablas

Tabla 6.1	Valores PCC y tiempos de procesamiento medio para la identificación del verdor en imágenes originales y tratadas con el <i>vignetting</i> y el balance de blancos.	154
Tabla 6.2	Valores PCC medios para la identificación del verdor en imágenes subexpuestas, normales (con y sin ajuste del TE) y sobreexpuestas; tiempo de procesamiento de cada método (μ s).	158
Tabla 6.3	Valor PCC y tiempo de procesamiento medio para la segmentación basada en $YD_B D_R$.	161
Tabla 6.4	Número de patrones procesados, vectores soporte obtenidos y parámetros obtenidos basados en el marco de los SVM.	164
Tabla 6.5	Valores de los parámetros para la fase de decisión durante la identificación de plantas.	165
Tabla 6.6	Porcentaje de acierto medio y tiempos de procesamiento (ms) de imágenes capturadas con la cámara SVS-VISTEK SVS4050CFLGEA para la detección de líneas de cultivo con diferentes alturas de plantas y rangos de búsqueda del ángulo de guiñada.	168
Tabla 6.7	Porcentaje de acierto medio y tiempos de procesamiento (ms) de imágenes capturadas con la cámara BASLER scA1400-17FC para SA, regresión lineal y Theil-Sen con diferentes reducciones de tamaño y márgenes de tolerancia.	174
Tabla 6.8	Porcentaje de acierto medio y tiempos de procesamiento (ms) de imágenes capturadas con la cámara SVS-VISTEK SVS4050CFLGEA para SA, regresión lineal y Theil-Sen con diferentes márgenes de tolerancia.	177

Lista de Abreviaturas

ADC	<i>Analog to Digital Converter</i> , convertidor analógico-digital.
AHRS	<i>Attitude and Heading Reference System</i> , Sistema de Referencia de Inclinación y Rumbo.
AP	Agricultura de Precisión.
APS	<i>Active Pixel Sensor</i> , Sensor de píxel activo.
BMP	<i>Bitmap format</i> , formato de imagen de mapa de bits.
CAR	Centro de Automática y Robótica.
CCD	<i>Charge Coupled Device</i> , Dispositivo de carga acoplada.
CMOS	<i>Complementary Metal–Oxide–Semiconductor</i> , semiconductor de óxido metálico complementario.
CSIC	Centro Superior de Investigaciones Científicas.
DLL	<i>Dynamic-link library</i> , Biblioteca de enlace dinámico.
DOP	<i>Dilution of Precision</i> , Dilución de la precisión.
EASA	<i>Environmentally Adaptive Segmentation Algorithm</i> , Algoritmo de segmentación adaptativo ambientalmente.
ECEF	<i>Earth-Centered Earth-Fixed</i> .
GPS	<i>Global Position System</i> , Sistema de Posicionamiento Global.
HSI	H, <i>hue</i> -tono; S, <i>saturation</i> -saturación; I, <i>intensity</i> -intensidad.
IMU	<i>Inertial Measurement Unit</i> , Unidad de Medición Inercial.

IP [code]	<i>International Protection</i> , Protección internacional.
IR	<i>Infrared</i> , infrarrojo.
ISCAR	Ingeniería de Sistemas, Control, Automatización y Robótica.
LIDAR	<i>Light Detection And Ranging</i> .
LUV	L, <i>luminance-luminosidad</i> ; U y V <i>chrominance-crominancia</i> .
LLH	<i>Latitude, Longitude, Height</i> ; Latitud, Longitud, Altura.
MLE	<i>Maximum Likelihood Estimation</i> , Estimación de máxima verosimilitud.
NDVI	<i>Normalized Difference Vegetation Index</i> , Índice de vegetación diferencial normalizado.
NED	<i>North-East-Down</i> , Norte-Este-Abajo.
PCC	Porcentaje de Clasificación Correcto.
RGB	R, <i>red-rojo</i> ; G, <i>green-verde</i> ; B, <i>blue-azul</i> .
RHEA	<i>Robot Fleets for Highly Effective Agriculture and Forestry Management</i> .
ROI	Region Of Interest, Región de interés.
SIG	Sistema de Información Geográfica.
SVI	<i>Space-Vehicle Information</i> , Información Espacial del Vehículo.
SVM	<i>Support Vector Machines</i> , Máquinas de Vectores Soporte.
UAV	Unmanned Aerial Vehicle, Vehículo Aéreo no Tripulado.
UPM	Universidad Politécnica de Madrid.
UTC	<i>Coordinated Universal Time</i> , Tiempo Universal Coordinado.
UTM	<i>Universal Transverse Mercator</i> , Sistema de coordenadas Universal Transversal de Mercator.
UV	<i>Ultraviolet</i> , ultravioleta.
Y D _B D _R	Y, <i>luminance-luminosidad</i> ; D _B y D _R <i>chrominance-crominancia</i> .
YUV	Y, <i>luminance-luminosidad</i> ; U y V <i>chrominance-crominancia</i> .

1

Introducción

1.1 Antecedentes

Las malas hierbas causan pérdidas muy importantes tanto cuantitativas como cualitativas en los cultivos de producción a nivel mundial. Se considera mala hierba a aquella planta que crece en un lugar no deseado y que además su presencia representa un perjuicio económico. Estas plantas crecen en medio de los cultivos y compiten con ellos por el agua del suelo, los nutrientes, el espacio y la luz, reduciendo su rendimiento y en ocasiones portando enfermedades que afectan al cultivo (Chris, 2012).

El control de malas hierbas es necesario debido a que las plantas que actúan como malas hierbas suelen ser muy competitivas, con una gran capacidad de adaptación a condiciones adversas y un rápido crecimiento. El tratamiento mediante agroquímicos es el método más extendido. Uno de los problemas graves de esta práctica es que algunas de esas malas hierbas llegan a hacerse resistentes a los herbicidas habituales, lo que ocasiona el uso de compuestos más caros y potentes aplicados en tasas más elevadas (GRDC, 2012). Dentro del tratamiento con agroquímicos, la práctica más utilizada en la agricultura tradicional es la de aplicar herbicida pulverizado de forma homogénea por todo el terreno, ignorando la variabilidad espacial y temporal debido al alto coste extra que supone un tratamiento desigual, ya que requiere un conocimiento preciso a priori de la distribución de la mala hierba en el cultivo. Las grandes dosis de herbicida y la resistencia de las malas hierbas constituyen un problema serio en la agricultura mundial, lo que provoca que el uso de agroquímicos sin control tenga

efectos negativos, entre los que destacan el gasto innecesario de herbicida (pérdida económica) y el daño medioambiental (contaminación del suelo y aguas subterráneas).

Una consecuencia directa de lo anterior es que se hace necesario el uso de sistemas de tratamiento específicos por zonas (Gerhards y Oebel, 2006), los cuales potencialmente mejoran la eficiencia en el control de malas hierbas y reducen las dosis de herbicida utilizado. El conjunto de técnicas orientadas a optimizar el uso de los insumos agrícolas en función de la variabilidad espacial (diferencias del cultivo en un mismo campo, en una misma campaña y cosecha) y temporal (cambios del cultivo en un mismo campo, en distintas campañas de cosecha) del cultivo es lo que se conoce como Agricultura de Precisión (AP). Así pues, se puede decir que la AP consiste en efectuar en el cultivo la intervención correcta, en el momento adecuado y en el lugar preciso.

La puesta en práctica de la AP implica la utilización de herramientas y tecnologías con diferente nivel de desarrollo, independientes entre sí pero completamente ligadas desde el punto de vista de consecución del objetivo final (Gil, 2001). Estos sistemas se pueden agrupar en cuatro niveles de tecnologías a utilizar:

- 1) Sistemas de Posicionamiento Global (GPS), los cuales permiten determinar la posición exacta de los equipos agrícolas durante el trabajo.
- 2) Captadores y sensores, materiales capaces de detectar, tratar y almacenar información sobre la variabilidad intra-parcelaria.
- 3) Sistemas de Información Geográfica (SIG), herramientas que permiten la gestión georreferenciada de la información parcelaria.
- 4) Tecnologías capaces de controlar automáticamente los equipos agrícolas, de forma que pueden modificar en continuo las condiciones de regulación de los mismos.

Del último punto se puede inferir que las nuevas tecnologías en AP, las cuales incluyen los sistemas de visión artificial, deben ser dispositivos que han de poder realizar cambios en continuo, sin que sea necesario detener la tarea que se esté realizando. Además, han de regularse en función de las condiciones que se presenten, como por ejemplo condiciones medioambientales adversas, y trabajar en conjunto con el resto de tecnologías.

Para los sistemas de visión artificial la precisión en la detección de malas hierbas y su distribución espacial constituyen un reto de interés, ya que ambas operaciones se tienen que realizar con una precisión que permita el tratamiento posterior de la mala hierba mediante sistemas de pulverización u otras técnicas de control entre las que se pueden encontrar el uso de quemadores, como se propone en el proyecto RHEA (2014) que se detalla más adelante.

Estos sistemas de visión artificial se pueden clasificar dependiendo de si el tratamiento se realiza en tiempo real o no:

- *Sistemas directos*: permiten trabajar a tiempo real realizando una actuación, sin la supervisión de un operario, sobre las distintas zonas a tratar inmediatamente después de la captura de la imagen y de la detección de malas hierbas.
- *Sistemas indirectos*: no permiten trabajar a tiempo real. Estos sistemas requieren de un procesamiento posterior de los datos obtenidos mediante la captura de imágenes, y por lo tanto no permiten una actuación inmediata a la detección sobre las zonas a tratar, siendo necesaria la supervisión por parte de un operario para la interpretación de los datos a la hora de identificar las zonas de mayor infesta. Esto a su vez origina una mayor compactación del terreno y mayor gasto de combustible al ser necesario realizar otro recorrido adicional sobre el cultivo a tratar.

En cuanto a los sistemas de visión artificial indirectos, hay que destacar el auge que están teniendo los *drones* dentro de la agricultura. Un *dron* o vehículo no tripulado, UAV (*Unmanned Aerial Vehicle*), no es más que una aeronave sin un piloto humano a bordo. Su vuelo se controla de forma autónoma mediante ordenadores o de forma remota mediante un operario situado en tierra u otro vehículo. El despegue y el aterrizaje también se pueden realizar de forma automática o tele operada. Existen distintos tipos de UAVs con distintos tamaños y configuraciones, entre los que destacan los aviones, helicópteros y multicopteros (helicópteros con varios rotores).

Si bien el uso de estos dispositivos en otros ámbitos, como puede ser el militar, los convierte en sistemas directos capaces de realizar escuchas o localización de objetivos, su uso en la agricultura los convierte en sistemas indirectos ya que no

permiten una actuación de forma directa y a tiempo real sobre el cultivo. Esto se debe en gran medida a las distintas limitaciones que tienen estos sistemas tales como su capacidad de vuelo limitada por el tipo de combustible o fuente de energía, el tamaño, el alcance o radio de actuación, el sistema de navegación, o su empuje, ya que motores más potentes pueden levantar mayores pesos pero incrementa el consumo y limita la altura que pueden alcanzar, siendo inviable a día de hoy el poder levantar grandes cargas como las que supondría un dispositivo pulverizador repleto de herbicida. Además, estos sistemas requieren que las señales emitidas y de control sean seguras, están muy limitados por el ancho de banda para transmitir datos y el espectro de frecuencias se encuentra muy saturado.

En definitiva, en la actualidad por lo general los sistemas de visión utilizados en aeronaves dentro del campo de la agricultura se centran en estimar y evaluar el estado de salud de la vegetación a partir de la medición de la radiación que las plantas emiten o reflejan, creando mapas que permiten detectar la actividad fotosintética discriminando así entre las distintas cubiertas vegetales, tal y como muestra la figura 1.1 (Peña y col., 2012, 2013).

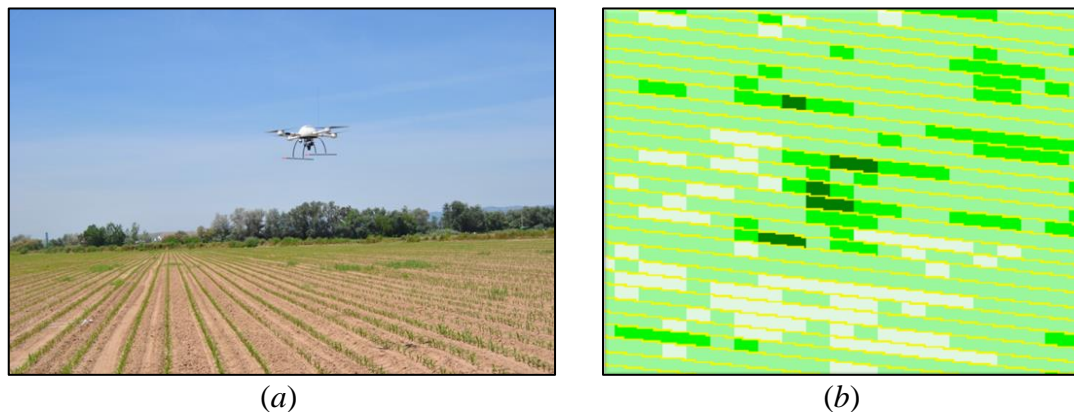


Figura 1.1 (a) UAV volando sobre un campo de maíz para detectar rodales de malas hierbas con el fin de generar un mapa de tratamientos específico por zonas. (b) Mapa de malas hierbas obtenido. (Imágenes cortesía de F. López-Granados y J.M. Peña; Instituto de Agricultura Sostenible, CSIC-Córdoba, España; Peña y col., 2012, 2013).

Por todo lo anterior, los sistemas de visión artificial indirectos en AP suponen una gran fuente de información pero no permiten una detección en tiempo real y actuación simultánea sobre el terreno. Un ejemplo de este tipo de sistema se puede encontrar en Zhang y col. (2010), quienes identifican la *Bidens Pilosa L* en campos de trigo a través de un sensor multimodal. Para ello, en colaboración con Eaton y col.

(2005) disponen de un robot (CASPA) que puede ser controlado remotamente o programado para una navegación autónoma mediante GPS. Este robot está equipado con un sistema de visión que permite captar imágenes en el modelo de color RGB e infrarrojo cercano.

En contraposición, los sistemas de visión artificial directos instalados sobre robots terrestres o tractores posibilitan una actuación inmediata sobre el campo. La mayoría de estos sistemas de visión artificial directos han sido desarrollados para detectar las plantas verdes en terrenos agrícolas mediante diferentes técnicas, a destacar las que usan imágenes RGB (espectro visible) y aquellas que utilizan el infrarrojo cercano. Un ejemplo de este sistema se puede encontrar en Berge y col. (2012), quienes mediante un robot equipado con un pulverizador y guiado por GPS, adquieren imágenes en formato RGB e identifican las malas hierbas en cereal utilizando una herramienta de software llamada Weedcer desarrollada por SINTEF (2014). Esta herramienta proporciona datos de estimación que facilitan la clasificación de los píxeles de la imagen dependiendo de si pertenecen al suelo o a una planta.

En base a las consideraciones anteriores, los requisitos necesarios de un sistema de visión artificial directo para trabajos agrícolas son los siguientes:

- a) Tener una alta resolución, de manera que se puedan detectar las malas hierbas y controlar el guiado del vehículo con la mayor precisión posible.
- b) Permitir realizar una localización espacial, de modo que se disponga de la distribución de malas hierbas georreferenciada y se pueda controlar el solapamiento de zonas tratadas evitando reiteraciones innecesarias.
- c) Ser completamente independiente y válido para cualquier sistema de tratamiento que pueda incorporar la visión como fuente de información.
- d) Actuar en tiempo real, de manera que el tratamiento se realice in-situ, sin necesidad de que las imágenes sean procesadas a posteriori y sin la supervisión de un operario.

1.2 Identificación del problema

Como se ha comentado anteriormente, a día de hoy se hace imprescindible el uso de sistemas de visión artificial para la identificación y localización espacial de malas hierbas en extensas superficies de cultivo con la finalidad de realizar una gestión de precisión asumible económicamente.

La evolución tanto de los tractores como de los aperos utilizados en el control de las malas hierbas implica que los operarios necesiten conocimientos muy específicos para llevar a cabo dicha tarea. Por ello, es necesario el uso de sistemas de visión que faciliten en gran medida las tareas a desarrollar, los cuales, en conjunto con otros sistemas mencionados en la sección 1.1 pueden hacer que tanto el control del tractor como los tratamientos agrícolas se realicen de manera autónoma, paliando en parte las restricciones horarias o temporales.

Durante las etapas iniciales en el desarrollo del maíz, al igual que en otros cultivos, la aparición de malas hierbas supone un grave perjuicio ya que éstas compiten con el maíz por los nutrientes, el agua y la luz. Por ello es importante el control de la mala hierba con el fin de evitar que en las zonas infestadas el crecimiento del maíz se vea perjudicado, siendo imprescindible la eliminación de las infestaciones a la vez que se preservan las plantas de maíz. El objetivo final es tener un cultivo libre de malas hierbas.

Como se verá detalladamente en la sección 1.3.1, existen dos proyectos principales (RHEA y AGL2011) que tienen como objetivo el desarrollo de tecnología para el control preciso de malas hierbas, lo que acarrea la identificación de malas hierbas existentes en el cultivo con el fin de aplicar tratamientos selectivos en función de las diferentes zonas e infestaciones existentes. Es importante destacar que la AP está presente en todo momento, desde la siembra del maíz hasta su recolección, por ello, cuando se habla de AP no se trata sólo de la identificación de malas hierbas y su tratamiento, sino que también es importante que la siembra se realice de forma precisa. El maíz se siembra creando líneas de plantas con una distancia entre líneas de 75 cm. Las malas hierbas pueden aparecer entre las líneas de cultivo o en la misma línea entre plantas. Dado que el sistema de visión artificial ha de trabajar a tiempo real, la

identificación de infestación entre plantas resulta más compleja y por tanto más costosa que la identificación entre líneas. Por otra parte, las malas hierbas más perjudiciales son aquellas que presentan patrones de crecimiento agregados, es decir que crecen formando rodales, por tanto en los casos más dañinos, una infestación que aparece entre líneas de cultivo también conlleva, por lo general, la existencia de mala hierba en las líneas. En este proceso, la identificación de líneas de cultivo con la mayor precisión posible es una tarea crucial, ya que permite identificar el verde existente entre dichas líneas para estimar la cantidad de mala hierba existente en zonas localizadas. Además, dichas líneas pueden servir como referencia para realizar el guiado automático del vehículo que está realizando el tratamiento y evitar que, en casos como el maíz, las ruedas pisen el cultivo.

En definitiva, el presente trabajo se centra en el desarrollo de métodos y procedimientos basados en un sistema de visión artificial directo para la identificación de líneas de cultivo en campos de maíz y densidades de vegetación existentes entre líneas, tanto para la aplicación de tratamientos selectivos como para un correcto guiado del vehículo. Concretamente el sistema físico de visión artificial se sitúa en un tractor autónomo, tal y como se muestra en la figura 1.2.

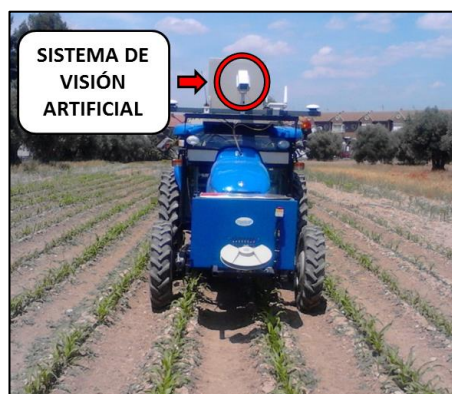


Figura 1.2 Sistema de visión artificial directo.

1.3 Motivación y objetivos

1.3.1 Motivación

El trabajo de investigación desarrollado tiene su origen en los proyectos de investigación llevados a cabo dentro del marco de investigación en el grupo ISCAR (2014), oficialmente reconocido por la Universidad Complutense de Madrid desde el año 2006.

El grupo ISCAR trabaja en diversos temas de investigación, siendo la robótica un área de especial interés y dentro de ésta la percepción computacional basada en sistemas de visión por computador uno de los aspectos que se abordan en profundidad. En relación a los sistemas de visión por computador se trabaja en los proyectos que se relacionan a continuación, algunos de ellos en colaboración con otros grupos de investigación tanto de carácter nacional como internacional. En estos proyectos existen vehículos autónomos dotados de equipos de visión para las misiones que se especifican. Una de las características más relevantes de los proyectos abordados es que los vehículos trabajan en entornos complejos, técnicamente conocidos como no estructurados donde la adaptación de los métodos tanto al entorno como a las condiciones impuestas por el propio movimiento de los vehículos, constituyen retos tecnológicos de envergadura, que se abordan en el proyecto de investigación que se presenta. A continuación se relacionan los proyectos que financian la investigación del trabajo de tesis descrito en esta memoria:

- 1) *Robot Fleets for Highly Effective Agriculture and Forestry Management* (RHEA), *THEME NMP-2009-3.4-1, Seventh Framework Programme*, (<http://www.rhea-project.eu/>) proyecto financiado por la Comisión Europea dentro del 7º Programa Marco (FP7) con duración del 1 de agosto de 2010 a 31 de julio de 2014. Este proyecto coordinado por el Centro de Automática y Robótica (CAR) -centro mixto del Consejo Superior de Investigaciones Científicas (CSIC) y la Universidad Politécnica de Madrid (UPM)- fue muy bien evaluado por la Comisión en su convocatoria. Está financiado con 6.59 M€ y en él participan diecinueve grupos pertenecientes a distintos organismos de investigación y empresas, entre ellas la Universidad Complutense a través del grupo ISCAR, que tiene asignadas las tareas relativas a la percepción visual de los vehículos terrestres. En este proyecto se investigan formas más inteligentes de proteger los cultivos, para ello, se dispone de una flota de robots aéreos (*drones*) y terrestres (tractores autónomos) interconectados con capacidades cartográficas y de teledetección avanzadas que permiten realizar en tiempo real tratamientos precisos para el control de malas hierbas mediante pulverización o por medios mecánicos y térmicos, o la fumigación de árboles, ya que cada tratamiento va asociado a un cultivo en concreto, siendo cereal, maíz y olivo respectivamente. Para cada cultivo se dispone de un tractor autónomo, diseñados

como prototipos, actuando en su conjunto como una flota de robots. Es en este proyecto en el que se centra la mayor parte de la investigación científica, concretamente en la detección y localización de líneas de cultivo y malas hierbas en campos de maíz mediante la aplicación de técnicas de visión por computador. La detección de las líneas es esencial, ya que permite tanto la identificación de las malas hierbas, por la existencia de vegetación alrededor de las mismas, como la detección de un elemento estructural de ayuda al guiado de los vehículos.

- 2) *Sistema de inspección terrestre en vehículos autónomos y su aplicación efectiva a la detección de malas hierbas y su control localizado* (AGL2011-30442-C02-02), perteneciente al Plan Nacional de I+D+i (01-01-2012 a 31-12-2014) proyecto coordinado por el Centro de Automática y Robótica (CAR) -centro mixto CSIC-UPM- y en el que participa la Universidad Complutense a través del grupo ISCAR. El objetivo final del proyecto es desarrollar e integrar tecnologías de detección aérea y terrestre y demostrar su validez en objetivos agronómicos de gran calado científico como es la discriminación entre plántulas de malas hierbas y cultivo. Para ello se utilizan imágenes de elevada resolución espacial provenientes de cámaras instaladas en vehículos aéreos no tripulados (1 – 5 cm pixel) y de sensores (≤ 1 cm pixel) situados en vehículos terrestres autónomos con el fin último de minimizar el uso de herbicidas. Los cultivos elegidos para la experimentación en campo son girasol y maíz. El criterio para la selección de estos cultivos se basa en su elevado impacto agro-económico en España y otros países así como la posible extrapolación de los resultados que se obtengan a otros cultivos igualmente importantes que se siembran en hilera ancha. El trabajo científico realizado en la presente tesis constituye parte del trabajo realizado en este proyecto.
- 3) *Sistema de vigilancia, búsqueda y rescate en el mar mediante colaboración de vehículos autónomos marinos y aéreos* (DPI2009-14552-C02-01), DPI (2009), ampliado hasta finales de 2013. Este proyecto incluye tareas basadas en visión para la localización de barcos y naufragos en el mar. Estas tareas requieren del desarrollo de técnicas y procedimientos basados en visión por computador en entornos sin estructurar donde las condiciones medioambientales son cambiantes.

Estos proyectos plantean verdaderos retos tecnológicos de futuro en las áreas especificadas, donde la visión computacional entraña su propia y compleja problemática, que constituye el objetivo de la investigación propuesta con el fin de que los vehículos involucrados consigan la máxima eficiencia sin pérdida de autonomía.

Los sistemas de visión en los entornos no estructurados, donde han de moverse los vehículos autónomos, se caracterizan por ofrecer soluciones ante condiciones altamente cambiantes y variables del entorno. Tal es el caso de la iluminación natural, donde ésta no sólo varía de un día a otro, sino que en cortos periodo de tiempo puede presentar fuertes variaciones. Pero no es éste el único problema, ya que los sistemas de visión se ubican en vehículos en movimiento, donde el propio movimiento (terrenos irregulares) o los efectos producidos por las vibraciones del motor pueden afectar considerablemente al sistema de visión, y en consecuencia, a la calidad de las imágenes capturadas, imposibilitando en algunos casos la identificación de los elementos de interés. Por este motivo, se plantea una investigación encaminada a lograr métodos de visión robustos que permitan obtener resultados satisfactorios en los escenarios descritos.

En RHEA y AGL2011, el problema se centra en la identificación de las líneas de cultivo en imágenes agrícolas, exactamente en cultivos de maíz, cuya finalidad estriba en la necesidad de discriminar entre lo que es propiamente cultivo frente a malas hierbas. Como se ha mencionado previamente, el objetivo final consiste en localizar los rodales de malas hierbas con el propósito de actuar sobre ellos con elementos mecánicos, químicos o térmicos para su total eliminación. En este contexto, la correcta localización de las líneas de cultivo constituye la clave del éxito para la detección y posterior eliminación de las malas hierbas, siendo también una pieza esencial en el control del guiado de los vehículos dentro del cultivo.

1.3.2 Objetivos

Los objetivos del presente proyecto de Tesis son los que se describen a continuación, todos ellos basados en los proyectos RHEA y AGL2011 antes mencionados, y derivados en gran medida de la problemática surgida tras la aplicación y puesta en marcha de los sistemas diseñados y desarrollados para estos proyectos:

- a) Estudio y análisis de técnicas de *reconocimiento* de elementos de interés basadas en técnicas de Visión por Computador, cuya finalidad es la identificación de los elementos y texturas de interés existentes en las imágenes procedentes de los entornos sobre los que navegan y actúan los vehículos autónomos. Destacan aquí la segmentación de las texturas relativas a la capa vegetal (cultivo y malas hierbas) y la identificación de las líneas de cultivo.
- b) Estudio y análisis de técnicas de *localización espacial* de los elementos de interés para la navegación autónoma y la actuación sobre ellos. Tras el reconocimiento es necesario proceder a su ubicación en el espacio en relación a un sistema de referencia asociado al sistema de visión.
- c) Investigación de la problemática real derivada de las condiciones adversas de navegación de los robots (vibraciones, irregularidades del terreno, condiciones climatológicas), así como de iluminación, todas ellas afectando a los métodos computacionales de tratamiento de imágenes de los sistemas de visión por computador.
- d) Propuesta de soluciones, mejoras y adaptación de los métodos investigados para su procesamiento en tiempo real.
- e) Análisis y validación de resultados de las soluciones propuestas.

1.4 Metodología

De cara a la consecución de los objetivos propuestos se realiza la planificación que se especifica junto con la estrategia a seguir, que constituye la metodología propuesta. Ésta se concreta en las tareas que se detallan a continuación:

Tarea 1 (objetivo *a*): estudio y análisis de métodos de técnicas de reconocimiento orientados a la identificación de texturas en Agricultura.

- Métodos de segmentación y reconocimiento de texturas con orientación específica en cultivos agrícolas con existencia de malas hierbas.
- Identificación de líneas de cultivo en campos de maíz.

Tarea 2 (objetivo *b*): estudio y análisis de métodos de localización espacial aplicando restricciones geométricas del sistema de visión.

- Geometría del sistema de visión.
- Ajuste preciso de las líneas de cultivo.

Tarea 3 (objetivo *c*): análisis de la problemática respecto de los métodos identificados previamente como los más prometedores.

- Identificación de texturas y líneas de cultivo.
- Métodos de localización espacial mediante técnicas específicas de posicionamiento.

Tarea 4 (objetivo *d*): propuesta de soluciones para aplicaciones en tiempo real considerando la problemática de los entornos de exterior y los efectos adversos provocados por los vehículos en movimiento.

- Búsqueda de nuevos métodos robustos de identificación de texturas y líneas de cultivo.
- Búsqueda de nuevos métodos de localización utilizando técnicas geométricas, teniendo en cuenta la problemática ya mencionada de los entornos de exterior y evitando el solapamiento de zonas tratadas.
- Búsqueda de estrategias robustas para contemplar el movimiento de los vehículos autónomos debido a las irregularidades del terreno, así como los cambios en la iluminación con alta tasa de variación.
- Búsqueda de estrategias para el control del guiado de los vehículos siguiendo las líneas de cultivo.
- Propuesta de un sistema robusto, independiente y adaptable, con un diseño abierto que permita la adaptación futura a otros cultivos.

Tarea 5 (objetivo *e*): análisis y validación de resultados.

- Validación de los métodos de identificación de texturas y líneas de cultivo.
- Validación de métodos relativos a la localización espacial.

- Análisis de la robustez de los métodos ante las condiciones adversas a las que se ven sometidos.
- Validación de los métodos de guiado.
- Verificación de los métodos de control de solapamiento de imágenes.

1.5 Aportes a la investigación

La contribución principal de este trabajo consiste en el desarrollo de un sistema de visión artificial enfocado a la identificación y localización de malas hierbas en campos de maíz, el control del guiado del vehículo y el control del solapamiento de las zonas tratadas. Este sistema constituye un hito indispensable en los proyectos RHEA y AGL2011. En concreto, en el proyecto RHEA el sistema diseñado y desarrollado se ha integrado en uno de los tractores de la flota, demostrando que el tractor era capaz de navegar siguiendo la línea de cultivo de maíz con suficiente precisión para no dañarla pisándola. Todos los datos y estudios llevados a cabo en el presente trabajo se han divulgado tanto en el ámbito nacional como internacional. A continuación se detallan los trabajos publicados que se apoyan en la investigación desarrollada mencionando la novedad aportada así como la relación con los temas abordados en el presente trabajo de tesis.

1.5.1 Publicaciones en Revistas Indexadas

1. **Guerrero, J.M.**, Pajares, G., Montalvo, M., Romeo, J., Guijarro, M. (2012a). Support Vector Machines for crop/weeds identification in maize fields. *Expert Systems with Applications*, vol. 39, issue 12, pages 11149-11155, doi: 10.1016/j.eswa.2012.03.040.

Se plantea una estrategia basada en aprendizaje para la identificación de plantas enmascaradas y sin enmascarar, que constituyen las texturas a detectar, basada en las Máquinas de Vectores Soporte. Para la fase de aprendizaje, primero se realiza una división de los píxeles en dos clases, posteriormente se obtienen los vectores soporte de cada clase y finalmente se calculan los valores medios de cada conjunto para establecer los límites de separación entre las clases. Ya en la fase de decisión, se valora cada píxel en función de dos reglas, las cuales determinan si pertenece a

una planta enmascarada o sin enmascarar. El método también es totalmente válido para evaluaciones post-tratamiento del campo. Esta estrategia aparece explicada en detalle en el capítulo cuatro sección 4.4.

2. **Guerrero, J.M.**, Guijarro, M., Montalvo, M., Romeo, J., Emmi, L., Ribeiro, A., Pajares, G. (2013). Automatic expert system based on images for accuracy crop row detection in maize fields. *Expert Systems with Applications*, vol. 40, issue 2, pages 656-664, doi: 10.1016/j.eswa.2012.07.073.

En este trabajo se plantea una técnica de identificación de líneas de cultivo basada en el estimador Theil-Sen. Para ello, conocidos los parámetros extrínsecos e intrínsecos del sistema de visión, se aplica una proyección de perspectiva para estimar las líneas de cultivo. Con las líneas estimadas y dado que a menudo la correspondencia con las reales no es precisa, se aplica una corrección. Esta corrección se lleva a cabo obteniendo los píxeles pertenecientes a verdes que están cercanos a las líneas estimadas y aplicando el conocido y robusto estimador Theil-Sen. Esta técnica aparece desarrollada en detalle en el capítulo cinco sección 5.4.3.

3. Montalvo, M., Pajares, G., **Guerrero, J.M.**, Romeo, J., Guijarro, M., Ribeiro, A., Ruz, J.J., de la Cruz, J.M. (2012a). Automatic detection of crop rows in maize fields with high weeds pressure. *Expert Systems with Applications*, vol. 39, issue 15, pages 11889-11897, doi: 10.1016/j.eswa.2012.02.117.

Se presenta una técnica de detección de líneas de cultivo en campos de maíz con una alta densidad de mala hierba basada en una doble umbralización. Después, con la ayuda de una plantilla se obtienen los píxeles que forman parte de las líneas de cultivo y se calcula la recta que pasa por dichos píxeles mediante mínimos cuadrados totales aplicando la técnica implementada para la referencia 2, lo cual constituye el aporte de este trabajo. Este método aparece descrito en el capítulo cinco sección 5.4.3.

4. Romeo, J., Pajares, G., Montalvo, M., **Guerrero, J.M.**, Guijarro, M., Ribeiro, A. (2012a). Crop Row Detection in Maize Fields Inspired on the Human Visual Perception. *The ScientificWorld Journal*, Volume 2012, Article ID 484390, 10 pages, doi:10.1100/2012/484390.

Trabajo que plantea una técnica para la identificación de líneas de cultivo en campos de maíz basada en la búsqueda de acumuladores máximos de píxeles verdes. Para ello, se realiza una búsqueda secuencial de las líneas candidatas en aquellas posiciones donde supuestamente se ubican las líneas de cultivo buscadas aplicando unos márgenes suficientemente amplios. Para la extracción de verdes se aplica una técnica basada en lógica borrosa que se describe en el capítulo cinco sección 5.3. En el presente trabajo se ha realiza un análisis comparativo de la técnica de extracción propuesta con el método de Otsu.

5. Pajares, G., **Guerrero, J.M.**, Romeo, J., Sánchez-Benítez, D., Montalvo, M., Herrera, P.J., Guijarro, M., de la Cruz, J.M., Ruz, J.J. (2012). Deterministic Simulated Annealing applied to Computer Vision tasks. Journal of High Performance Computing. Bioinfo Publications, Volume 2, Issue 1, pp.-20-23, (<http://bioinfopublication.org/viewhtml.php?artid=BIA0000418>). ISSN: 2230-7192 & E-ISSN: 2230-7206.

Trabajo que lleva a cabo una revisión sobre la aplicación del enfoque de optimización de enfriamiento simulado (*simulated annealing*) para resolver tres problemas diferentes de visión por computador en entornos de exterior como es el caso del trabajo desarrollado en la presente memoria de tesis: detección de cambios en la imagen, correspondencia estereoscópica y clasificación de texturas. En él se aplica la técnica que se descrita en la referencia 9.

6. Montalvo, M., **Guerrero, J.M.**, Romeo, J., Emmi, L., Guijarro, M., Pajares, G. (2013a). Automatic expert system for weeds/crops identification in images from maize fields. Expert Systems with Applications, vol. 40, issue 1, pages 75-82, doi: 10.1016/j.eswa.2012.07.034.

En este trabajo se plantea el uso de una doble umbralización mediante el método de Otsu para la identificación de plantas enmascaradas y sin enmascarar. El aporte realizado consiste en comparar el método propuesto con el planteado en la referencia 1. Esta técnica aparece descrita en el capítulo cuatro sección 4.4.4.

7. Romeo, J., Pajares, G., Montalvo, M., **Guerrero, J.M.**, Guijarro, M., de la Cruz, J.M. (2013a). A new Expert System for greenness identification in agricultural images. *Expert Systems with Applications*, Volume 40, Issue 6, Pages 2275-2286 doi: 10.1016/j.eswa.2012.10.033.

Se propone un sistema experto para la identificación del verde en imágenes agrícolas. La estrategia se divide en dos módulos, el primero, encargado de la toma de decisión basándose en el análisis del histograma de la imagen, y el segundo, encargado de la identificación del verde aplicando dos estrategias distintas. La primera está basada en los métodos clásicos de identificación del verde, que es la aportación de esta tesis, aplicando las técnicas estudiadas en la referencia 1 a este trabajo, y la segunda inspirada en un agrupamiento borroso (*fuzzy clustering*).

8. Romeo, J., **Guerrero, J.M.**, Montalvo, M., Emmi, L., Guijarro, M., Gonzalez de Santos, P., Pajares, G. (2013b). Camera sensor arrangement for crop/weeds detection accuracy in agronomic images. *Sensors*, 13, 4348-4366, doi: 10.3390/s130404348.

Este trabajo constituye una ampliación del presentado en la referencia 11, que aparece posteriormente. En él, se identifican los parámetros extrínsecos e intrínsecos más relevantes del sistema de visión con el fin de obtener las imágenes con la mayor resolución posible. Una vez identificados se calculan los valores óptimos de forma empírica.

1.5.2 Publicaciones en Revistas no Indexadas

9. **Guerrero, J.M.**, Pajares, G., Guijarro, M. (2011). Técnicas de procesamiento de imágenes estereoscópicas. Enlaces: revista del CES Felipe II, ISSN 1695-8543, N°.13

Se trata de una revisión de métodos para correspondencia estereoscópica con el fin de encontrar el mejor de los posibles y además mejorarlo. Se aprovechan las bondades de los algoritmos estudiados para diseño y desarrollo del nuevo método. Se utilizan técnicas de suavizado comunes tales como la media o la mediana junto con un método de enfriamiento simulado y otro de propagación que se ha creado

expresamente para este fin. Los resultados de estas dos técnicas se comparan con los métodos clásicos previamente seleccionados.

1.5.3 Congresos y conferencias

10. **Guerrero, J.M.**, Montalvo, M., Romeo, J., Oliva, D., Guijarro, M., Pajares, G. (2012*b*). Detection of images with high CCD charge level due to indirect sunlight in maize fields. International Conference of Agricultural Engineering, julio 8-12, Valencia (España). ISBN-10: 84-615-9928-4 / ISBN-13: 978-84-615-9928-8 (<http://cigr.ageng2012.org>).

El objetivo final de esta estrategia es determinar si una imagen dada es válida para la identificación de cubierta vegetal o es inapropiada debido a los altos valores de intensidad producidos por un exceso de iluminación. Esta técnica aparece descrita en detalle en el capítulo cuatro sección 4.2.1.

11. **Guerrero, J.M.**, Romeo, J., Emmi, L., Montalvo, M., Guijarro, M., Pajares, G., González de Santos, P. (2012*c*). Influence of the vision system pitch angle on crop and weeds detection accuracy. First RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture, septiembre 19-21, Pisa (Italia). ISBN-13: 978-88-6741-021-7 (<http://www.rhea-conference.eu/2012/>).

En este trabajo se realiza un estudio en campos de maíz simulados mediante la herramienta Webots (2014) sobre la influencia del ángulo de cabeceo (*pitch*) del sistema de visión cuando se mantienen constantes el resto de parámetros del sistema (ángulos de alabeo, guiñada y altura). En él se analiza la resolución de la imagen y la precisión. Este trabajo es además el precursor del trabajo presentado en la referencia 8.

12. **Guerrero, J.M.**, Montalvo, M., Guijarro, M., Romeo, J., Pajares, G. (2014). Comparison of two new automatic techniques for exposure time control in agricultural images. Second RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture and forestry, mayo 21-23, Madrid (España) (<http://www.rhea-conference.eu/2014/>).

Las cámaras utilizadas en sistemas agrícolas han de hacer frente a los diferentes cambios de iluminación provocados por trabajar en entornos de exterior. En este trabajo se presentan dos técnicas automáticas para controlar el tiempo de exposición. Uno de los métodos, descrito en el capítulo seis sección 6.3.3, está basado en una simplificación del enfoque propuesto en la referencia 17 y utiliza una imagen de referencia para el ajuste. La otra, se centra en el estudio del canal de intensidad obtenido por la transformación de la imagen del modelo RGB al HSI y utiliza el tiempo de exposición actual para llevar a cabo el ajuste. Esta última técnica aparece descrita en el capítulo cuatro sección 4.2.2.2 apartado *b*.

13. Montalvo, M., **Guerrero, J.M.**, Romeo, J., Herrera, P.J., Guijarro, M., Ribeiro, A., Pajares, G. (2011). Camera System geometry for site specific treatment in precision agriculture. In Proc. First RHEA Workshop (G. Rabatel and P. González de Santos), September 9, Montpellier (France).

En este trabajo se realiza un estudio comparativo entre dos tipos de cámara con el fin de verificar sus comportamientos con diferentes parámetros extrínsecos e intrínsecos. Mediante este estudio es posible identificar la mejor configuración geométrica que permite obtener una imagen con la mayor resolución posible. Para este trabajo se ha aplicado el modelo geométrico descrito en la referencia 2 y que se explica en detalle en el capítulo cinco sección 5.4.1.

14. Montalvo, M., **Guerrero, J.M.**, Romeo, J., Oliva, D., Guijarro, M., Pajares, G. (2012*b*). Unsupervised learning for crop/weeds discrimination in maize fields with high weeds densities. International Conference of Agricultural Engineering, julio 8-12, Valencia (España). ISBN-10: 84-615-9928-4 / ISBN-13: 978-84-615-9928-8 (<http://cigr.ageng2012.org>).

Se plantea una técnica de aprendizaje no supervisado basado en un aprendizaje supervisado por cuantificación vectorial (LVQ, Learning Vector Quantization). Mediante la aplicación de un doble umbral a través del método de Otsu se consigue que el enfoque sea no supervisado. El aporte principal del presente trabajo lo constituye la identificación del verde mediante la aplicación del índice vegetativo ExG.

15. Romeo, J., Oliva, D., **Guerrero, J.M.**, Montalvo, M., Guijarro, M., Pajares, G. (2012*b*). Plants discrimination in maize fields applying human perception to digital images. International Conference of Agricultural Engineering, julio 8-12, Valencia (España). ISBN-10: 84-615-9928-4 / ISBN-13: 978-84-615-9928-8 (<http://cigr.ageng2012.org>).

Se propone un método para cuantificar la diferencia cualitativa entre dos colores. Para ello se aplica una modificación de la distancia de Manhattan que corrige los errores de ésta cuando la diferencia entre dos colores no es una diferencia equilibrada en los tres canales espectrales. La aportación principal en este trabajo lo constituye la implementación del algoritmo en MATLAB y sus resultados preliminares.

16. Guijarro, M., **Guerrero, J.M.**, Montalvo, M., Romeo, J., Pajares, G. (2013). A new approach to solve image thresholding in precision agriculture. 9th European Conference on Precision Agriculture. Lleida, Catalonia, Spain, July 7-11, poster. (<http://www.ecpa2013.udl.cat/>).

Se propone un estudio comparativo de distintos métodos de binarización para la identificación del verdor en imágenes agrícolas. El aporte principal de este trabajo es la implementación de los distintos métodos de umbralización.

17. Montalvo, M., **Guerrero, J.M.**, Romeo, J., Guijarro, M., de la Cruz, J.M., Pajares, G. (2013*b*). Acquisition of Agronomic Images with Sufficient Quality by Automatic Exposure Time Control and Histogram Matching. Lecture Notes in Computer Science. Advanced Concepts for Intelligent Vision Systems (ACIVS'13) (Jacques Blanc-Talon, Andrzej Kasinski, Wilfried Philips, Dan Popescu, Paul Scheunders, Eds.), 8192, 37-48. Springer International Publishing Switzerland. October 28-31, Poznan, Poland.

En este trabajo se presenta un procedimiento para la adquisición de imágenes con suficiente calidad controlando el tiempo de exposición. Esto se lleva a cabo mediante el estudio del histograma de la imagen y se completa con una correspondencia de histogramas. Además, se realiza un estudio aplicando la conversión del modelo RGB al HSI, lo que constituye la aportación principal de

este trabajo. Una variación de esta técnica se aplica y compara con la estrategia propuesta en la referencia 12.

18. Montalvo, M., **Guerrero, J.M.**, Guijarro, M., Romeo, J., Pajares, G. (2014). Exposure time control in agricultural images using a color adaptive method. Second RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture and forestry, mayo 21-23, Madrid (España) (<http://www.rhea-conference.eu/2014/>).

En este trabajo se propone un nuevo método para calcular el tiempo de exposición de la cámara para la adquisición de imágenes agrícolas. Este método ajusta el tiempo de exposición en función de la imagen anterior, que además se estudia para decidir si el ajuste ha de ser fuerte o débil en función del rango de variación en la iluminación. Para la obtención de los resultados es necesario el procesamiento de las imágenes para su posterior evaluación, esto constituye el aporte principal de este trabajo.

19. Romeo, J., Andrea Benítez, **Guerrero, J.M.**, Montalvo, M., Pajares, G. (2014). Tractor guidance in maize fields using computer vision and geometry. Second RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture and forestry, mayo 21-23, Madrid (España) (<http://www.rhea-conference.eu/2014/>).

Se presenta un algoritmo basado en la visión por computador para guiar un tractor autónomo a lo largo de las líneas de cultivo en campos de maíz. La obtención de imágenes de calidad es indispensable para el guiado preciso, por lo que para la obtención de las imágenes con el sistema de visión disponible se ha utilizado el método propuesto en la referencia 12.

1.5.4 Seminarios

20. Conferencia en la Facultad de Informática de la Universidad Complutense de Madrid (mayo 2011), titulada “Sistemas de visión estereoscópica para control de robots de exterior y en agricultura de precisión”, en el marco de convenio de colaboración del Fondo de Cooperación Internacional en Ciencia y Tecnología Unión Europea-México (FONCICYT). Con la participación del Instituto

Politécnico Nacional, la Universidad de Guadalajara (México), Universidad Libre de Berlín y Universidad Complutense de Madrid. En esta conferencia se comenzaron a establecer las problemáticas existentes entre las imágenes y su captura y calidad en los entornos de exterior.

1.5.5 Emisiones radiofónicas

21. Participación en el programa de radio “SIN DISTANCIAS” titulado “Ojos artificiales en agricultura” emitido el 21 de marzo 2012 dentro de la programación de la Universidad Nacional de Educación a Distancia (UNED) en RADIO 3 de RNE, relativo a la utilización de vehículos robotizados en entornos agrícolas equipados con sistemas de visión.

1.6 Organización de la tesis

La memoria de tesis se presenta estructurada en capítulos. Estos capítulos se organizan siguiendo el orden natural de la investigación, cuya distribución es la que se ofrece a continuación:

Primer capítulo. *Introducción*, donde se estudian y plantean los objetivos de la tesis, así como la motivación e introducción de las propuestas y planteamientos realizados, tal y como se ha expuesto en las secciones precedentes.

Segundo capítulo. *Estado del arte*, donde se revisan las diferentes estrategias existentes en la literatura que proporcionan los argumentos y la base de la investigación.

Tercer capítulo. *Procesamiento de imágenes en agricultura de precisión*, donde se describen, desde un punto de vista teórico, los métodos necesarios para el tratamiento de imágenes agrícolas y que forman parte de la investigación realizada en esta tesis. También sirve como introducción a los sucesivos capítulos donde se facilitan nuevas alternativas a las técnicas ya existentes.

Cuarto capítulo. *Segmentación de imágenes agrícolas en campos de maíz*, donde se plantean nuevas técnicas que permiten el control de la iluminación y el color para la obtención de imágenes de calidad, permitiendo realizar una separación entre la cubierta vegetal y el resto de elementos de manera eficiente. En este capítulo, se plantea un nuevo método de detección y control de imágenes con alta intensidad luminosa

mediante el tiempo de exposición, así como distintas formas de tratamiento tales como el balance de blancos o la corrección del efecto *vignetting*. También, se propone un nuevo método de segmentación y umbralización basado en el modelo de color $YD_B D_R$, y una nueva técnica basada en Máquinas de Vectores Soporte que permite la distinción entre plantas enmascaradas y sin enmascarar.

Quinto capítulo. *Sistema experto: detección de líneas de cultivo y malas hierbas, guiado y solapamiento*, donde se presenta la arquitectura de un sistema abierto que permite la identificación de las líneas de cultivo basándose en la geometría del sistema de visión. Una vez detectadas las líneas, éstas son corregidas mediante el estimador Theil-Sen para alcanzar una mayor precisión, ya que dichas líneas sirven para: detectar la cobertura de malas hierbas existentes en la zona de estudio, realizar un control del guiado del vehículo, y controlar del solapamiento de las zonas ya tratadas. Este sistema consta del sistema físico de visión artificial (cámara, lente y filtros) en conjunto con una IMU dotada de GPS, todo ello encapsulado en una carcasa con protección IP66 para trabajos de exterior.

Sexto capítulo. *Análisis de resultados*, donde se presentan los resultados obtenidos con las nuevas técnicas propuestas en los capítulos precedentes y se analizan comparativamente con otros métodos existentes en la literatura. El estudio comparativo se ha realizado tanto de manera cuantitativa como cualitativa, siendo un aspecto importante el criterio del experto humano.

Séptimo capítulo. *Conclusiones y trabajo futuro*, donde se exponen las líneas generales seguidas, realizando un balance general del trabajo. Asimismo, se plantean las líneas de investigación futuras así como posibles mejoras o extensiones de los métodos propuestos.

Octavo capítulo. *English summary*, donde se realiza un resumen en inglés de las principales aportaciones de este trabajo de tesis.

2

Estado del arte

2.1 Visión general

Desde el inicio de la Robótica y de la Inteligencia Artificial, el ser humano ha querido dotar a las máquinas de capacidades similares a los seres vivos. Debido a esto, son numerosas las líneas de investigación que se han abierto con el fin de dotar a las máquinas de sentido visual, y es que gracias a este sentido, una máquina con percepción visual es capaz de afrontar problemas que sin ella quedarían sin resolver. En su mayoría, los sistemas con la capacidad sensorial de la visión se complementan con otros mecanismos sensoriales tales como sensores de alcance o proximidad (Fu, 1988) para después completar el proceso realizando una integración multisensorial (Lopez-Orozco, 1999). En la presente tesis, el centro de estudio se basa en la identificación de líneas de cultivo y malas hierbas en campos de maíz utilizando únicamente técnicas de visión computacional. Como ya se mencionó en la introducción, en el capítulo cinco sección 5.4.2 se propone el uso de una Unidad de Medición Inercial (*Inertial Measurement Unit*, IMU en terminología inglesa) como complemento que permite conocer la posición exacta del sistema de visión, en relación a un sistema de coordenadas virtual absoluto, y así llevar a cabo el proceso de identificación de líneas de cultivo y mala hierba con la mayor precisión posible.

Desde mediados de los años 60, el procesamiento de imágenes ha crecido enormemente con diferentes finalidades, aunque todas ellas tienen unas tareas comunes como son la de realzar y extraer la información contenida en las imágenes para su

posterior interpretación. Por ello, Pajares y Cruz (2007) describen la visión artificial por computador como la capacidad de una máquina para ver el mundo que le rodea, deduciendo la estructura y las propiedades del mundo tridimensional a partir de una o más imágenes bidimensionales.

En definitiva, todo proceso basado en visión computacional, conlleva implícitamente la definición de un sistema de visión basado en una o más cámaras con sus correspondientes sistemas ópticos, capaces de capturar imágenes para su posterior procesamiento, en este caso mediante técnicas propias de inteligencia artificial. Estos dos aspectos son los que se abordan en el presente capítulo con la perspectiva de su aplicación al tema objeto de la investigación desarrollada, que como se ha indicado anteriormente es la identificación de líneas de cultivo en campos de maíz junto con las malas hierbas existentes para su posterior tratamiento selectivo, así como el guiado lo más preciso posible del vehículo siguiendo las líneas de maíz sin dañarlas. En las secciones 2.2 y 2.3 se analizan los aspectos relacionados con las cámaras, el sistema óptico y la formación de imágenes. En las secciones 2.4 y 2.5 se revisan los métodos existentes en la literatura que tratan una problemática similar a la abordada en esta tesis.

2.2 Cámaras

Existen diferentes dispositivos para la captura de imágenes (Ballard y Brown, 1982), pero el elemento principal para la adquisición de imágenes es la cámara, siendo objeto de este estudio las cámaras digitales. La parte principal de una cámara digital es su sensor. Este sensor no es más que un chip formado por miles o millones de componentes de distintos materiales sensibles a la luz, los cuales al estar expuestos a la radiación y la luminosidad procedentes de los objetos en la escena, son capaces de capturar este tipo de información que se traduce en lo que se conoce como imagen digital.

Los sensores más utilizados son de dos tipos: CCD (*Charge Coupled Device*, dispositivo de carga acoplada) y CMOS (*Complementary Metal–Oxide–Semiconductor*).

2.2.1 Sensor CCD

Este sensor es uno de los más utilizados en la formación de la imagen digital. Proporciona buena calidad de imagen, pero su fabricación es muy compleja y costosa.

Las cámaras digitales que llevan incorporado esta clase de sensor, tienen un coste de compra elevado y su consumo de energía es relativamente superior. En parte, este sobrecoste se debe a que para el funcionamiento del sensor CCD se necesita de un chip externo denominado ADC (*Analog Digital Converter*, convertidor analógico-digital en español), encargado de convertir los datos de cada píxel en datos digitales binarios, para que otro dispositivo los pueda leer.

El funcionamiento es muy simple, el sensor CCD se encarga de transformar las cargas de las celdas de la matriz en voltajes y de entregar una señal analógica de salida que será posteriormente digitalizada por la electrónica asociada a la cámara, es decir, se hace una lectura de cada uno de los valores correspondientes a las celdas y esta información se traduce a datos mediante un convertidor analógico-digital. La lectura de estas cargas se realiza mediante desplazamientos sucesivos y de forma secuencial. La figura 2.1 muestra un esquema genérico de un sensor CCD.

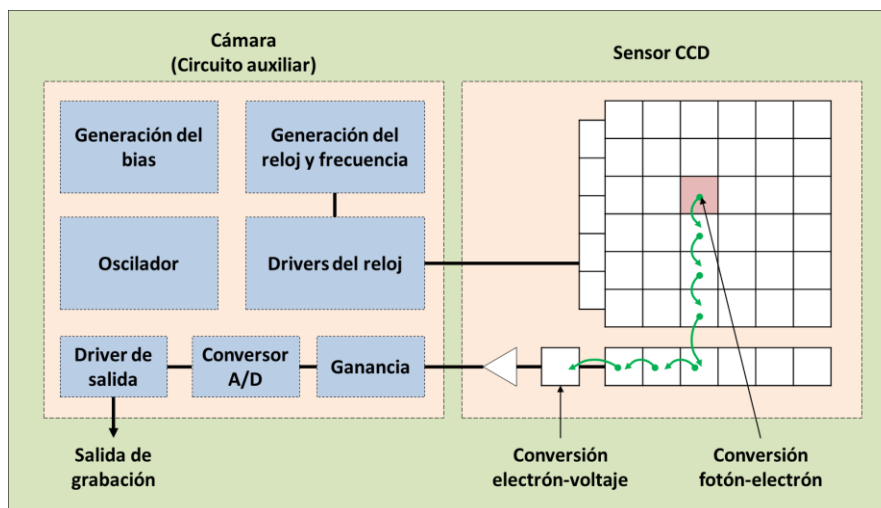


Figura 2.1 Esquema de un sensor CCD.

2.2.2 Sensor CMOS

En un sensor CMOS (también llamado APS, *Active Pixel Sensor*), cada celda es independiente del resto. La principal característica es que la digitalización de los píxeles se realiza de forma interna mediante una serie de transistores acoplados a cada celda, por lo que todo el trabajo se lleva a cabo dentro del sensor y no se necesita un chip externo, lo cual se traduce en una reducción de costes a la vez que permite el diseño de equipos más pequeños. La figura 2.2 muestra un esquema de la disposición de los distintos elementos que componen un sensor de tipo CMOS.

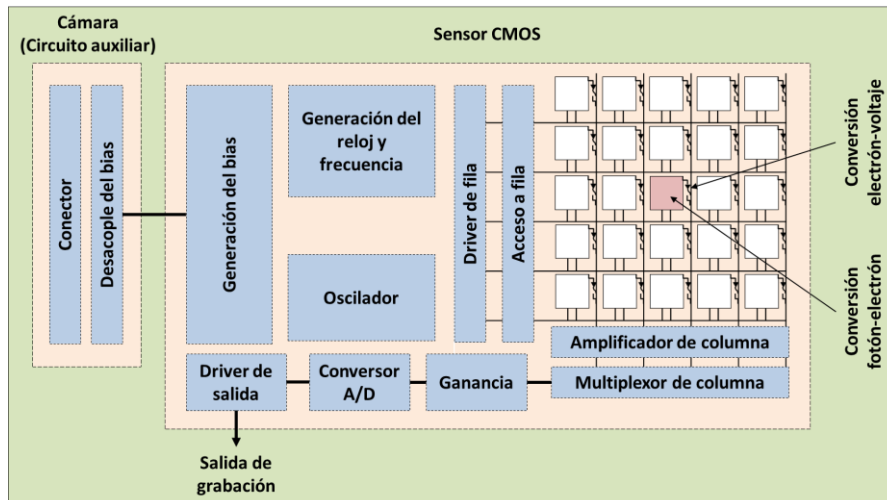


Figura 2.2 Esquema de un sensor CMOS.

2.2.3 Comparativa de sensores CCD y CMOS

Existen diferentes estudios donde se comparan los resultados obtenidos mediante estos dos tipos de sensores, por ejemplo, en Helmers y Schellenberg (2003) se realiza una comparación mediante su aplicación al campo de la interferometría, mientras que en Magnan (2003) la comparación es según su arquitectura y proceso de fabricación.

Por ello, a la hora de comparar los dos tipos de sensores, existen diferentes factores a considerar (Litwiller, 2001; Rashidian y Fox, 2011). Si se desea que con poca luz se obtenga una señal aceptable, los CMOS son superiores a los CCD debido a su mayor capacidad de respuesta (nivel de señal ofrecida por energía de óptica incidente) debido a la sencillez de integrar elementos amplificadores. Si por el contrario se desea un rango dinámico, determinado por el nivel de saturación de los píxeles y un umbral por debajo del cual no captan señal, los CCD superan a los CMOS en prestaciones. Otro factor importante es el ruido, donde los CCD obtienen mejores resultados debido a que por construcción todo el procesado se realiza fuera del sensor, pudiéndose aprovechar convertidores más rápidos o mejores procesadores, mientras que en los CMOS, al realizar la mayoría de las funciones dentro del sensor, pierden espacio para el fotodiodo en sí; si bien, debido a que la conversión analógico-digital se realiza en el propio sensor CMOS, la velocidad de éstos es bastante superior a los CCD. Relacionado con el ruido está la respuesta uniforme, mediante la cual se espera que un píxel ante el mismo nivel de excitación no presente cambios apreciables respecto a sus vecinos. Por lo tanto, dado

que en los CMOS los píxeles son individuales, éstos son más propensos a sufrir más fallos que los CCD, lo que se traduce en un aumento del nivel de ruido. Otra característica de la individualidad de los CMOS es que permiten seleccionar porciones de la imagen más fácilmente y por lo tanto, tomar imágenes a menor resolución sin necesidad de realizar operaciones de interpolación. Y por último, cabe destacar el llamado efecto *blooming*, consistente en la propagación del nivel de saturación entre píxeles vecinos cuando un píxel se ha saturado, siendo un fenómeno que ocurre en los CCD y que por construcción no ocurre en los CMOS. Este fenómeno origina la aparición de zonas altamente brillantes en las imágenes debido a la propagación de la saturación en regiones de la imagen más o menos extensas. La figura 2.3 muestra dos dispositivos del tipo CCD y CMOS.

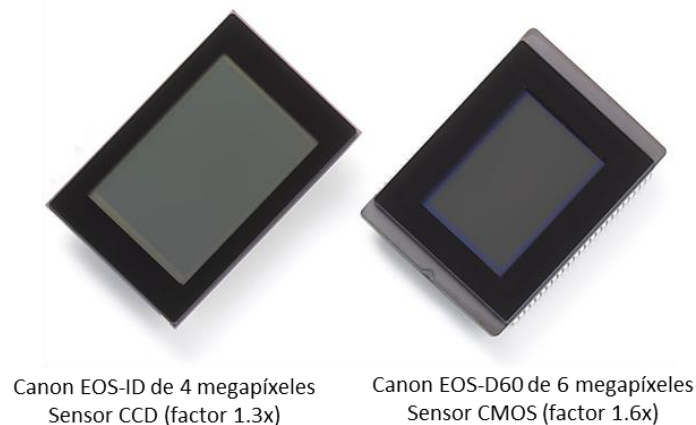


Figura 2.3 Ejemplos de sensores con sus respectivos factores de multiplicación de la distancia focal, CDD a la izquierda y CMOS a la derecha.

A modo de resumen, se puede concluir que un sensor CCD obtiene mayor calidad de imagen a costa de un tamaño mayor y un consumo más elevado, pero como siempre, no hay nada superior de por sí, sino situaciones en las que cada tecnología resulta más o menos adecuada en función de la aplicación concreta.

La investigación desarrollada y reflejada en la presente memoria, se refiere al tratamiento de imágenes en campos de cultivo agrícolas, y por tanto se trata de entornos de exterior, sometidos a condiciones de iluminación extremas y altamente variables.

Bajo esta perspectiva, las características que ha de tener un sistema de visión autónomo para aplicaciones de exterior y en particular en campos agrícolas pueden sintetizarse en las siguientes:

- a) Amplio rango dinámico para ajustar el nivel de carga que recibe el sensor, de este modo es posible ajustar la cantidad de luz que recibe el sensor en función de las condiciones ambientales variables y en ocasiones adversas.
- b) Gran calidad de imagen con el menor ruido posible, ya que se requiere una identificación lo más precisa posible de las plantas, tanto de cultivo como de malas hierbas, que constituyen uno de los elementos esenciales a tratar en las imágenes.
- c) Una respuesta uniforme donde los píxeles vecinos no difieran en exceso entre sí, con el fin de realizar un procesamiento de la imagen a nivel de vecindad entre píxeles en contraposición al procesamiento a nivel de regiones.

Teniendo en cuenta las características anteriores, el sensor CCD supera en este caso al CMOS, por lo que las cámaras elegidas para la realización de las pruebas llevadas a cabo durante la investigación desarrollada y descritas en el capítulo seis utilizan un sensor de tecnología CCD.

2.2.4 Formación de la imagen a color

Para la formación de una imagen a color no basta sólo con un sensor, ya que éste es monocromo y sólo obtiene un valor proporcional a la intensidad de la luz y la radiación recibida, por lo que no tiene color. Para captar la imagen a color es necesario disponer de un sistema de filtros que cubran dicho sensor, de forma que cada píxel, en el sensor, va a ser más o menos sensible a la radiación recibida en función de las longitudes de onda en el espectro visible y por tanto a la naturaleza de dichas longitudes correspondientes al rojo, verde y azul.

Uno de estos filtros, el más utilizado, es el conocido como mosaico Bayer, y que se caracteriza porque su diseño se ha desarrollado teniendo en cuenta la capacidad del ojo de ser más sensible a la luz verde que al resto. Los científicos consideran que este hecho se debe a la propia evolución del hombre en el hábitat natural, particularmente en las etapas iniciales de la evolución, donde el hábitat lo constituía la naturaleza en la que predominaban los verdes. Por ello, en el diseño del mosaico de Bayer se utiliza el doble de diodos verdes que de rojo y azul. Para cada píxel existe un filtro que limita la incidencia de la radiación a una única longitud de onda (rojo, verde o azul). Por lo tanto,

para la formación de la imagen se recorre todo el mosaico en grupos de cuatro píxeles, que aportan los datos de color a cada porción de imagen, y cuya transformación da lugar a una imagen en el espacio de color RGB, hablándose en este caso de los canales R, G y B.

La figura 2.4 muestra un esquema del filtro de Bayer, que permite la obtención de los tres canales R, G y B, observándose cómo el número de píxeles que permiten la incidencia de longitudes de onda correspondientes a la componente verde duplica al número de píxeles correspondientes a las componentes rojo y azul.

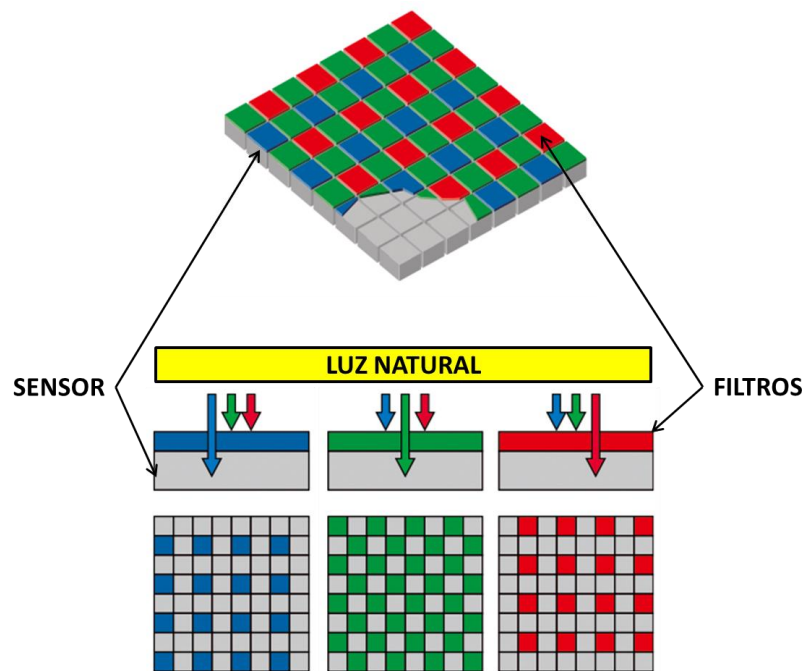


Figura 2.4 Esquema del mosaico Bayer.

2.3 Sistema óptico

Los sensores descritos previamente, suelen venir montados sobre un cuerpo que los protege y que dispone de las conexiones necesarias para su funcionamiento. La mayoría de las cámaras industriales utilizadas en los sistemas de visión artificial no disponen de un sistema óptico fijado al cuerpo, sino que permite la colocación de distintos sistemas para su adaptación a la aplicación donde se vaya a utilizar el sistema. Dos elementos esenciales a la hora de elegir un sistema óptico son la lente y el iris.

2.3.1 Distancia focal de la lente

La distancia focal es la distancia existente entre el centro óptico (ubicado en el centro de la lente) y la superficie sensible a la luz (sensor). La distancia focal se relaciona con el ángulo de visión en una relación inversa. Mientras más pequeña sea esta distancia, más grande será el ángulo de visión. Así, por ejemplo, un objetivo de 20 mm ofrece un ángulo de visión más grande que el que ofrece un objetivo de 200 mm. También se relaciona con el concepto de profundidad de campo (distancia que un objeto en la escena puede acercarse o alejarse de la lente manteniéndose enfocado), mientras mayor sea la distancia focal (mayor zoom), menor es la profundidad de campo y mayor es la sensación de proximidad del objeto.

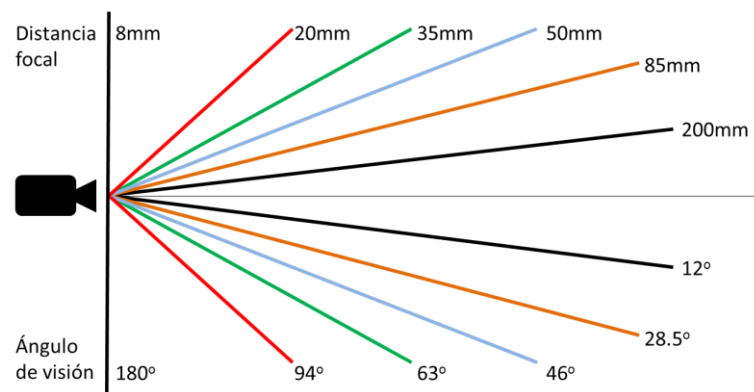


Figura 2.5 Relación entre distancias focales y ángulos de visión.

Gran parte de las lentes que se utilizan en sistemas de visión, son lentes con una distancia focal fija, que se deben usar con sensores del mismo tamaño (formato), o con de sensor de tamaño menor. No se deben utilizar con tamaños mayores de sensor (Kruegle, 2006) porque producen distorsiones inaceptables y un oscurecimiento en los bordes de la imagen (efecto *vignetting*, ver capítulo tres sección 3.2.2). Por ejemplo, cuando un fabricante indica que una lente es para un sensor de 1/3 pulgadas, ésta se puede utilizar en un sensor de 1/4 pulgadas pero no en uno de 1/2 pulgadas sin que se produzcan los efectos adversos mencionados. La figura 2.6 muestra una imagen obtenida con una lente de 1/2 de pulgada con un sensor de 2/3 pulgada. Se observa claramente un elevado efecto de distorsión radial, así como la inutilización de parte del sensor.



Figura 2.6 Ejemplo de sensor de 2/3" y lente de 1/2".

2.3.2 Iris: fijo, manual y automático

Una imagen deseable ha de ser nítida, clara, con una exposición correcta y un contraste y una resolución aceptable para la finalidad específica. Por ello, el iris de un sistema óptico basado en lentes constituye una pieza clave del mismo, ya que es el encargado de mantener un nivel de iluminación óptimo incidiendo sobre el sensor del sistema de visión, además de permitir el control correcto de la profundidad de campo.

a) *Iris de tipo fijo y manual*: a pesar de existir sistemas ópticos de iris fijo, éstos son sólo convenientes en ubicaciones de interior, ya que en ellos se puede controlar los niveles de iluminación, manteniéndolos más o menos constantes. También existen sistemas ópticos de iris manual, que permiten regular la cantidad de luz entrante girando unos anillos concéntricos para abrir o cerrar el iris y que se mantienen fijos durante el funcionamiento. Con estos dispositivos no es necesario cambiar el sistema óptico aunque el sistema de visión cambie su emplazamiento, y en ese caso sólo es necesario realizar un nuevo ajuste. Al igual que en el caso de los sistemas de iris fijo, no se recomienda su uso en entornos de luz variables como es el caso de las aplicaciones de exterior, salvo que la cámara permita el ajuste dinámico del tiempo de exposición en función de la iluminación entrante. Este último es el tipo de dispositivo utilizado en el trabajo de investigación descrito en esta memoria y que se detalla en profundidad en el capítulo cuatro sección 4.2.2.2.

Los sensores CCD y CMOS pueden capturar imágenes eficientemente, aún con grandes variaciones en el nivel de luz utilizando un iris manual, pero necesitan un sistema óptico de iris automático cuando se quiere que trabajen dentro de un amplio rango de iluminación, en otras palabras, desde una luz solar intensa hasta una situación de escasa iluminación como la nocturna (Kruegle, 2006).

b) *Iris de tipo automático*: este tipo de iris tiene un mecanismo electrónico por el cual la cantidad de luz que pasa a través de la lente se ajusta dependiendo de la intensidad luminosa en la escena y de la sensibilidad de la cámara. Existen dos tipos de iris automático, el DC y el de vídeo. Ambos suelen utilizar una señal analógica para el control de su apertura, y difieren en que los primeros integran el motor que lo regula dentro de la propia cámara mientras que en el segundo caso el motor se sitúa dentro del propio sistema óptico.

Las lentes tienen el inconveniente de que se pueden ver afectadas por la difracción y el desenfoque cuando la apertura del iris es demasiado pequeña, es por ello que existen otro tipo de sistemas ópticos cuyo iris se ha mejorado para solventar este problema, ya que limitan el cierre del iris. Uno de estos tipos de iris es el fabricado por la empresa Axis Communications (2014) denominado iris de tipo P, que ofrece mejoras en contraste, claridad, resolución y profundidad de campo, ya que dispone de un software asociado que optimiza la calidad de la imagen a partir del análisis del histograma de la imagen entrante. Tener una mejora en la profundidad de campo implica que objetos a distancias diferentes puedan ser enfocados de forma simultánea.

El inconveniente de los sistemas ópticos de iris automático es que son más caros que los manuales, y en el caso de los DC, la cámara va ligada al sistema óptico ya que necesita del motor que regula la apertura. La figura 2.7 muestra dos ejemplos de sistemas ópticos de iris manual (izquierda) y de iris automático (derecha). La figura 2.8 muestra la cámara con el CCD a la vista, así como el sistema óptico de iris manual con enfoque ajustable. También aparece un filtro de corte de las radiaciones infrarroja y ultravioleta, su acoplamiento se realiza en la parte anterior del sistema óptico. Este tipo de filtros resulta necesario para evitar que las radiaciones infrarrojas, muy abundantes en los entornos de exterior, impacten en el CCD, produciendo imágenes con los colores falseados. Algunos sistemas de visión acoplan este filtro cubriendo exactamente el propio sensor CCD o CMOS. El sistema mostrado en la figura 2.8 es el que se ha utilizado en el proyecto RHEA.



Figura 2.7 Ejemplos de sistemas ópticos, a la izquierda sistema con iris manual y a la derecha sistema con iris automático.

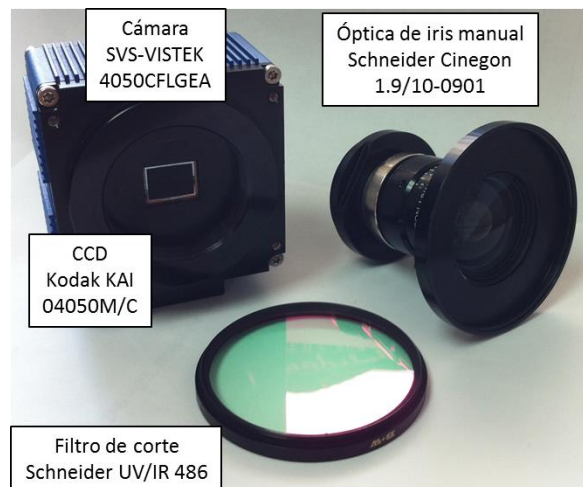


Figura 2.8 Cámara con sensor CCD, sistema óptico con iris manual y filtro de corte ultravioleta e infrarrojo.

2.3.3 Enfoque

El enfoque está ligado a la nitidez con la que se visualizan los objetos a una determinada distancia y puede ser fijo o automático. Son numerosos los campos en los que se utiliza un enfoque automático (Chang y col., 2009; Chen y col., 2010; Lee y col., 2011; Antti y col., 2012), pero en todos ellos se necesita un mecanismo de control que lo regule. Es por ello, que para la aplicación que se plantea en la presente tesis, donde la zona de estudio se sitúa siempre a una distancia fija, el enfoque fijo es una opción viable.

2.4 Técnicas de segmentación

En las secciones anteriores se han descrito tanto los componentes imprescindibles que forman parte de un sistema de visión como las características principales de éstos, con ello se han introducido los elementos hardware necesarios para obtener una imagen.

Tan importante como el hardware, lo es el software, mediante el cual se llevará a cabo el procesamiento de las imágenes para la identificación de líneas de cultivo y malas hierbas, el guiado del tractor autónomo y el control del solapamiento de las zonas de estudio. Todo esto se verá en detalle en los capítulos cuatro y cinco.

2.4.1 Métodos de segmentación

Segmentar una imagen digital significa dividirla en zonas disjuntas e individualizadas (Marcos, 2006). El objetivo de la segmentación es simplificar y/o cambiar la representación de una imagen en otra más significativa y más fácil de analizar. Visto de otra forma, se puede definir la segmentación como el proceso de asignar a cada píxel de una imagen una etiqueta que lo defina, de manera que la agrupación de puntos con la misma etiqueta y conectados espacialmente representan un objeto. Como norma general, existen dos tipos de enfoques en los métodos de segmentación:

1) Métodos basados en las características de los píxeles (Puig y García, 2006; Rud y col., 2006; Hanmandlu y col., 2004; Yu y col., 2002; Kumar y col., 2001).

2) Métodos basados en regiones (Puig y García, 2006; Randen y Husøy, 1999; Wagner, 1999; Smith y Burns, 1997).

Respecto a las diferencias, los métodos basados en píxeles estudian las propiedades específicas e individuales para cada píxel, mientras que los métodos basados en regiones tienen en cuenta la distribución espacial de los píxeles dentro de la textura y por lo tanto su relación con otros píxeles de la región.

Dependiendo de la técnica utilizada, estos enfoques pueden clasificarse según el método de extracción de características aplicado. Por un lado están las técnicas de agrupamiento borroso (*fuzzy clustering*). Dentro de estas técnicas se encuentran trabajos tales como el de Mohaddamzadeh y Bourbakis (1997) donde se segmentan agrupaciones de píxeles con un color similar; el de Cheng y col. (2001) donde tratan de identificar regiones de similares características definiendo el concepto de *homogeneidad fuzzy*; o los de Seong y Usery (2001) y Zhang y Foody (1998) quienes aplican las técnicas anteriores en el ámbito de las imágenes de satélite. Además, se pueden encontrar otros trabajos basados en agrupamiento borroso tales como Pajares y Cruz (2002), Pajares y col. (2002) y Lim y Lee (1990).

Por otro lado, se sitúan las técnicas basadas en redes neuronales. Trabajos tales como los de Mas y Flores (2008) se aplican en la clasificación de texturas en imágenes por satélite. En Huang (1999) se utilizan tres redes de tipo Hopfield (Haykin, 1994), una por cada canal de color, donde los nodos son los píxeles del correspondiente canal y las conexiones representan las similitudes entre canales. En Campadelli y col. (1997) se utiliza una red neuronal que requiere de un proceso de clasificación previo. Otros trabajos tales como Pajares y col. (2001), Atkinson y Tatmall (1997) o Paola y Schowengerdt (1995) utilizan una red neuronal de retro-propagación, mientras que los mapas Auto-Organizativos (SOM, *Self Organizing feature Maps*) se aplican también en trabajos tales como el de Jiji y Ganesan (2008) o el de Lo y Pei (1999).

En trabajos, como en los propuestos por Giacinto y col. (2000), Chan y col. (2003) o Del-Frate y col. (2007) se utiliza otra técnica basada en aprendizaje llamada cuantificación vectorial (*Learning Vector Quantization*), la cual puede ser supervisada proporcionando ejemplos de las distintas clases existentes en la imagen, o no supervisada en cuyo caso es el propio método el que selecciona el número de clases en función de un umbral previamente fijado. También se puede encontrar trabajos basados en la transformada de Fourier (Pajares y Cruz, 2007a; Gonzalez y Woods, 2007) tales como Lillo y col. (2007), Aiyer y col. (2005) o Pyun y col. (2002).

Otros métodos de segmentación pueden encontrarse en Frizelle y Moody (2001), Newsam y col. (2004), Xie y Tsui (2004), Trauwaert y col. (1991) y Benediktsson y col. (1990), todos ellos están basados en estimaciones de máxima verosimilitud, donde el aprendizaje reside en la evaluación de una función de densidad de probabilidad. Haapanen y col. (2004), Muñoz-Marí y col. (2007) y Singh y col (2001) utilizan el método k -NN (*k-Nearest Neighbours*) donde los píxeles se clasifican en función de la clase a la que pertenecen un determinado número de k vecinos próximos al píxel que se desea clasificar o segmentar. Tso y Mather (1999) basan el proceso de segmentación en algoritmos genéticos. Tottrup (2007), Kumar y Majumder (2001), Hansen y col. (1996) y Debeir y col. (2002) presentan una alternativa a los métodos clásicos basándose en reglas mediante la creación de árboles de decisión.

2.4.2 Revisión de métodos de segmentación en imágenes naturales

El aumento en el desarrollo de robots equipados con sensores de visión artificial aplicados en AP requiere de soluciones para diferentes problemas. El robot ha de navegar y actuar sobre una zona determinada dentro de un campo de cultivo de amplias dimensiones (Davies y col., 1998) donde gran parte de la información la proporciona el sistema de visión tanto para navegación como para detección de elementos de interés en la imagen (Rovira-Mas y col., 2010), por lo tanto, la segmentación de la imagen es una tarea importante relacionada con la aplicación de sistemas de visión artificial en AP.

La segmentación eficiente y automática de las imágenes para extraer la vegetación constituye un paso importante para muchas aplicaciones, tales como la identificación de malas hierbas para un tratamiento específico por zonas (Burgos-Artizzu y col., 2009, 2010; Guijarro y col., 2011; Onyango y Marchant, 2003; Tellaeche, 2008a, b, 2011). Además, la clasificación del suelo tiene una gran relevancia ecológica (Luscier y col., 2006). Por todo esto, existen numerosas estrategias propuestas con el fin de segmentar la capa vegetal en imágenes de cultivo que pueden agruparse en las tres categorías siguientes:

a) Estrategias basadas en índices del espectro visible: estas técnicas engloban los métodos de identificación que utilizan el índice de exceso de verde ExG (*Excess Green*, Ribeiro y col., 2005; Woebbecke y col., 1995), el índice de exceso de rojo ExR (*Excess Red*, Meyer y col., 1998), la extracción del índice de color de la vegetación CIVE (*Color Index of Vegetation Extraction*, Kataoka, 2003) y el exceso de verde menos el exceso de rojo ExGR (Neto, 2004). También forma parte de este conjunto el índice de vegetación (VEG) descrito en Hague y col. (2006), el cual se diseñó para hacer frente a la variabilidad en la iluminación procedente de la luz natural. Los índices ExG, ExGR, CIVE y VEG han sido aplicados de forma conjunta en Guijarro y col. (2011) mejorando el rendimiento con respecto a su aplicación individual, tal y como se verá en el capítulo tres sección 3.3. Todos estos enfoques necesitan de un umbral para realizar la segmentación final.

b) Estrategias basadas en una umbralización específica: estas técnicas incluyen una umbralización dinámica y generalmente asumen que se trata de un problema bi-clase (con dos clases o conjuntos), donde se han de identificar las plantas y

el resto (sombras, suelo, paja, etc.) como elementos constitutivos de dichas clases. Reid y Searcy (1987) estiman una función de decisión bajo el supuesto de que la distribución de las muestras en las clases siguen una distribución Gaussiana. El método de Otsu (1979) también se ha aplicado considerando un problema de dos clases (Ling y Ruzhitsky, 1996; Shrestha y col., 2004), todas estas aproximaciones se aplican a imágenes en escala de grises. Gebhardt y col. (2006) proponen una umbralización para segmentar la imagen transformándola desde el modelo de color RGB a una única imagen en escala de grises, algoritmo que más tarde fue mejorado aplicando una homogeneidad local y operaciones morfológicas en Gebhardt y Häußbauch (2006). Kirk y col. (2009) aplican una combinación del verde y de la intensidad derivada de las bandas espectrales roja y verde, y calculan un umbral automático para un problema entre dos clases asumiendo dos funciones de distribución de probabilidad Gaussianas, las cuales están asociadas al suelo y a la vegetación respectivamente y definen la distribución de las muestras en clases. Este procedimiento requiere de una estimación previa del ángulo de rotación del eje hipotético asociado al color verde. En Meyer y Carmago-Neto (2008) se utiliza una umbralización automática mediante el método de Otsu para binarizar una imagen en escala de grises obtenida mediante ExG y el índice de diferencia normalizado (NDI, *Normalized Difference Index*). En el trabajo se compara la segmentación obtenida con el método propuesto con la segmentación obtenida mediante ExGR. Se determina que en este último caso, un valor de cero es suficiente para el umbral, de modo que no es necesario utilizar el método de Otsu. Por otro lado, Guijarro y col. (2011) y Burgos-Artizzu y col. (2011) han aplicado el valor de la media estadística de la imagen transformada, la obtenida mediante los índices de vegetación, en vez de aplicar una umbralización automática. Esta elección se justifica en que el método de Otsu proporciona un valor de umbral mayor que la media y por lo tanto produce una infra-segmentación, que se traduce en la desaparición de partes de vegetación relevantes existentes en las imágenes, algo a todas luces no deseable.

c) Estrategias basadas en aprendizaje: dentro de este conjunto se sitúan las técnicas basadas en el agrupamiento borroso (o *fuzzy clustering*). Por ejemplo, Meyer y col. (2004) han aplicado enfoques no supervisados, incluyendo técnicas de agrupamiento borroso para la segmentación de diferentes zonas de interés a partir de los índices ExR y ExG. Existen otros tipos de técnicas tales como las descritas en Tian y Slaughter (1998), donde se propone el algoritmo de segmentación adaptativo

ambientalmente (EASA, *Environmentally Adaptive Segmentation Algorithm*) para la detección de plantas a través de un proceso de aprendizaje supervisado; Ruiz-Ruiz y col. (2009) aplican el algoritmo EASA bajo el espacio de color HSI para tratar la variabilidad en la iluminación; Zheng y col. (2009, 2010) utilizan un algoritmo supervisado, *Mean-Shift*, suponiendo que la separación entre las muestras asociadas a la vegetación y al suelo puede ser tratada como un problema de segmentación entre dos clases. La separabilidad de las clases se valida mediante una red neuronal de retro-propagación y el discriminante lineal de Fisher (1936) respectivamente; los espacios de color utilizados son RGB, LUV y HSI.

Como puede verse, la mayoría de los métodos aplican enfoques basados en las características espectrales de los píxeles ya que la identificación de texturas en entornos de exterior, donde la variabilidad meteorológica es un factor importante, impide que los métodos basados en regiones trabajen correctamente. Todos estos métodos se centran en la identificación del verde bajo el supuesto de que la componente espectral verde predomina sobre la componente espectral roja. Pero tal y como se verá en el capítulo cuatro, no siempre ocurre que ambas componentes sean claramente diferenciables. Por ello, se propone una estrategia basada en las Máquinas de Vectores Soporte (*Support Vector Machines*, SVM) (Guerrero y col., 2012a) con el fin de identificar las plantas cuya componente espectral verde se encuentra enmascarada por varias razones, y por lo tanto, su cercanía a la componente roja es mayor, llegando a veces a ser indistinguible. Esta técnica resulta ser una combinación entre las técnicas de aprendizaje y las que utilizan una umbralización específica, ya que se trata de resolver también un problema bi-clase. Tellaeché y col. (2011) aplican SVM para detectar cuándo una zona del campo ha de ser tratada, llevando a cabo la segmentación mediante el método descrito en Ribeiro y col. (2005). Rumpf y col. (2012) utilizan SVM de forma secuencial para discriminar, primero entre grupos de plantas de similares especies identificando correctamente entre monocotiledóneas, dicotiledóneas y cereal (cebada), y a continuación aplican SVM dos veces más para extraer las características que permitan distinguir entre las distintas plantas dicotiledóneas. Además, dentro del ámbito de las Máquinas de Vectores Soporte se encuentran trabajos tales como los descritos en Camps-Valls y Bruzzone (2005), Camps-Valls y col. (2004), Melgani y Bruzzone (2004) o Ferreiro-Armán y col. (2007), aunque ninguno de ellos aplicado a la identificación de la capa vegetal en imágenes naturales.

2.4.3 Umbralización de imágenes

La umbralización es el método más simple para la segmentación de una imagen. A partir de una imagen en escala de grises se puede utilizar la umbralización para crear imágenes binarias (Shapiro y Stockman, 2001). Un umbral sencillo basado en el valor medio del nivel de gris de la imagen (histograma) se ha propuesto en Gée y col. (2008), tras su aplicación, la parte vegetal (cultivo y mala hierba) aparece en blanco, y el resto (superficie del suelo, piedras, sombras) como negro. En Guijarro y col. (2011) se aplica el conocido método de Otsu, tradicionalmente utilizado en binarización de imágenes. Además, también se han aplicado otras estrategias más complejas como la utilizada en Bossu y col. (2009), basada en el método *k-means clustering* (agrupamiento *k-means*). En esta línea existen numerosos trabajos tales como los descritos en Sahoo y col. (1988), Gonzales-Barron y Butler (2006), basados en el de Weszka (1978), y Fu y Mu (1981) donde se estudian varios algoritmos de umbralización y se comparan las diferentes técnicas estudiadas en ellos. Tomando como referencia estos trabajos se puede realizar una clasificación de los diferentes métodos de umbralización según sean locales o globales, estando divididos estos últimos en métodos basados en píxeles o en regiones al igual que ocurre con los métodos de segmentación estudiados en la sección 2.4.1.

a) Umbralización global: dentro del conjunto de métodos que se basan en el estudio individual de cada punto o píxel de la imagen, las técnicas de umbralización global más destacadas son: P-tile (Doyle, 1962), método de modo (Prewitt y Mendelsohn, 1966), método de Otsu (1979), concavidad del histograma (Rosenfeld y De La Torre, 1983), entropía (Pun, 1980, 1981; Johannsen y Bille, 1982; Kapur y col., 1985), preservación del momento (Tsai, 1985), ISODATA (Ridler y Calvard, 1978) y mínimo error (Kittler y Illingworth, 1986).

Respecto a los métodos de umbralización global basados en regiones, es decir donde se tiene en cuenta la vecindad de cada punto así como su localización espacial, se encuentran los métodos basados en la transformación del histograma (Katz, 1965; Mason y col., 1975; Weszka y Rosenfeld, 1979, 1974; Weszka y col., 1974, 1973; Watanabe y grupo CYBEST, 1974; Wu y Rosenfeld, 1982), los métodos basados en estadísticas de segundo orden del nivel de gris (Haralick y col., 1973; Ahuja y Rosenfeld, 1978, Kirby y Rosenfeld, 1979), el método de Deravi y Pal (1983), los métodos de relajación (Southwell, 1940, 1946; Rosenfeld y Smith, 1981; Fekete y col.,

1981; Zucker y col., 1977; Pavlidis, 1977; Peleg, 1980), los métodos de relajación del gradiente (Bhanu y Faugeras, 1982), así como otros métodos más generales (Bartz, 1969; Wolfe, 1969; Ullman, 1974; Morrin, 1974; Panda, 1977).

b) Umbralización local: dentro de la umbralización local se encuentran trabajos tales como los de Chow y Kaneko (1971, 1972), Nakagawa y Rosenfeld (1979) y Fernando y Monro (1982).

Además, existen métodos específicos para la obtención de múltiples umbrales como son la segmentación de la amplitud (Boukharouba y col., 1985), el método de Wang y Haralick (1984) o el contraste uniforme (Kohler, 1981). Asimismo, muchos de los métodos mencionados anteriormente, como por ejemplo Otsu (1979), la entropía (Pun, 1980, 1981; Kapur y col., 1985), el algoritmo de preservación del momento (Tsai, 1985) o el de mínimo error (Kittler y Illingworth, 1986), pueden extenderse para la obtención de varios umbrales.

2.5 Técnicas de detección de líneas de cultivo y malas hierbas

Como se ha comentado en el apartado anterior, los sistemas de visión artificial instalados sobre robots autónomos suscitan un gran interés dentro de la agricultura por su aplicación a tratamientos específicos por zonas en tareas propias de AP. Por ello, los sistemas de visión suministran gran cantidad de información para permitir que el robot navegue y actúe sobre un área específica.

Una tarea importante relacionada con la aplicación de métodos de visión artificial es la que tiene que ver con la detección de líneas de cultivo y malas hierbas, y que ha dado lugar a numerosos estudios en este área (Burgos-Artizzu y col., 2009, 2010; Guerrero y col., 2012a; Lopez-Granados, 2011; Montalvo y col., 2012a; Onyango y Marchant, 2003; Sainz-Costa y col., 2011; Tellaeche y col., 2008a, b, 2011).

Para la detección de líneas de cultivo se han propuesto numerosas estrategias. Por ejemplo, Fontaine y Crowe (2006) pusieron a prueba las habilidades de cuatro algoritmos de detección de líneas para determinar la posición y el ángulo de la cámara respecto a un conjunto artificial de líneas de cultivo con y sin malas hierbas simuladas. Los métodos contrastados fueron el análisis de líneas (*stripe analysis*), la transformada de Hough, el reconocimiento de regiones (*blob analysis*) y la regresión lineal.

A continuación, se realiza un estudio de las diferentes técnicas de identificación de líneas, agrupadas en categorías que asimismo incluyen los métodos anteriormente mencionados.

2.5.1 Métodos basados en la exploración de líneas horizontales

Søgaard y Olsen (2003) aplican la transformación de una imagen RGB a una imagen en escala de grises. Esto lo realizan dividiendo la imagen de color en sus tres canales rojo, verde y azul y aplicando los métodos para la extracción de la capa vegetal descritos en Woebbecke y col. (1995), obteniendo una imagen en escala de grises que se divide en líneas horizontales donde los valores máximos indican la presencia de una línea candidata. Cada máximo determina un segmento lineal y el centro de gravedad del segmento se marca en esa posición de la línea. Las líneas de cultivo se identifican uniendo los puntos marcados a través de un método similar a la transformada de Hough o bien aplicando regresión lineal.

Sainz-Costa y col. (2011) han desarrollado una estrategia basada en el análisis de secuencias de video para la identificación de líneas de cultivo. Las líneas de cultivo persisten a lo largo de las direcciones definidas por la proyección en perspectiva con respecto a una escena tridimensional en el campo. Basándose en este hecho, aplican una transformación de la imagen a escala de grises utilizando la estrategia propuesta por Ribero y col. (2005) para posteriormente realizar una binarización de la imagen utilizando una técnica de umbralización. Cada imagen se divide en cuatro líneas horizontales. Se dibujan áreas rectangulares sobre la imagen binaria para identificar las zonas de cultivo y de mala hierba. Los centros de gravedad de esas áreas se utilizan como los puntos que definen la línea de cultivo y con los cuales se construye la ecuación de la recta. La primera captura es utilizada como una tabla de contenido (*lookup table*) que guía todo el proceso para determinar las posiciones donde las siguientes zonas pueden ser identificadas en posteriores capturas.

Hague y col. (2006) transforman la imagen original RGB en una imagen en escala de grises. La imagen transformada se divide en ocho franjas horizontales. La intensidad de los píxeles a través de esas franjas muestra una variación periódica debido a las líneas de cultivo paralelas. Dado que las características de la cámara, posición y espaciado de las líneas de cultivo se conocen a priori, el espaciado a nivel de píxel de

cada línea se puede calcular para cada franja horizontal utilizando el modelo de óptica de cámara estenopeica. Además, se puede construir un filtro de tipo paso banda con el fin de identificar el patrón buscado mediante una respuesta dada en el dominio de la frecuencia. El problema de estas técnicas es que a veces los patrones horizontales son difíciles de extraer ya que los cultivos y las malas hierbas forman una zona única.

2.5.2 Métodos basados en la transformada de Hough

De acuerdo a Slaughter y col. (2008), uno de los métodos más utilizados en la visión por computador para la identificación de líneas de cultivo se basa en la transformada de Hough (1962). Este método fue pensado para hacer frente al problema de aparición de líneas discontinuas y por tanto con huecos donde el cultivo no ha crecido debido a una mala germinación o a otros factores que dan lugar a la pérdida de plantas en la línea, como pueden ser una destrucción de las mismas por aplastamiento, un fallo en la siembra o a la acción de roedores y aves que utilizan las semillas sembradas como alimento. Esta técnica se ha utilizado también para el guiado automático de vehículos agrícolas en tiempo real (Astrand y Baerveldt, 2005; Hague y col., 1997; Leemans y Destain, 2006; Marchant, 1996). La transformada de Hough se aplica sobre imágenes binarias, las cuales se obtienen aplicando técnicas similares a las explicadas en la sección 2.4.2 y 2.4.3, es decir, transformaciones de una imagen RGB a escala de grises seguidas de una binarización (Tellaache y col., 2008*a, b*, 2011).

Gée y col. (2008) aplican una doble transformada de Hough bajo la hipótesis de que las líneas de cultivo son las únicas líneas existentes en la imagen que convergen hacia el conocido punto de fuga, mientras que las líneas restantes son descartadas; además, añaden restricciones adicionales tales como el espaciado entre líneas de cultivo y la geometría del sistema óptico basada en proyección de perspectiva, conceptos que ayudan a identificar las líneas. Este método requiere establecer el umbral necesario en la transformada de Hough para determinar el número máximo de valores pico que dicha transformada proporciona (Jones y col., 2009*a, b*) o los picos más relevantes (Rovira-Más y col., 2005). Dependiendo de la densidad del cultivo pueden aparecer varias líneas candidatas, por ello, para deshacer la ambigüedad se aplica posteriormente un proceso que une las líneas con parámetros similares (Tellaache y col., 2008*a, b*, 2011). Este método tiene la desventaja de su elevado coste computacional (Ji y Qi, 2011) que lo convierte en inadecuado para ser utilizado en un sistema que requiera

respuesta en tiempo real tal como el que se plantea en este trabajo de tesis. Por otro lado, la aleatoriedad de la transformada de Hough necesita que se seleccionen pares de puntos para que sean considerados como una línea, es decir, pares de puntos pertenecientes a líneas de cultivo. Si esta técnica se aplica en imágenes donde se han extraído los puntos de borde, la selección de esos pares llega a ser altamente compleja debido a que las malas hierbas forman parte del conjunto de bordes extraídos.

2.5.3 Métodos basados en el punto de fuga

Pla y col. (1997) proponen una técnica que identifica regiones (cultivo/malas hierbas y suelo) aplicando la segmentación en imágenes a color y determinando la orientación hacia el punto de fuga de las líneas paralelas en el cultivo. Por otro lado, Pitas (2000) y Komala y Punitha, (2009a, b, 2008) utilizan el esqueleto de cada región definida como una característica para encontrar las líneas que definen las líneas de cultivo. Los esqueletos obtenidos y sus propiedades, definidas como una sucesión de puntos del contorno conectados, permiten identificar las líneas de cultivo que están orientadas hacia el punto de fuga. Este proceso es muy dependiente de los esqueletos, los cuales no siempre son fáciles de extraer, especialmente teniendo en cuenta que existen rodales de mala hierba que se solapan con los cultivos.

En Romeo y col. (2012a) se aplica también conocimiento relativo a la posición del punto de fuga para detectar las líneas de cultivo. El proceso se basa en la identificación de acumuladores máximos de píxeles verdes pertenecientes a las líneas que están orientadas hacia el punto de fuga. En él se propone una técnica de agrupamiento borroso (*fuzzy clustering*) para la identificación del verde. Esto hace que el método dependa en gran medida de la fase de entrenamiento a diferencia de la técnica que se presenta en el capítulo cinco sección 5.4, que es no supervisada.

2.5.4 Métodos basados en visión estéreo

Kise y col. (2005) o Kise y Zhang (2008) desarrollaron un sistema de guiado para maquinaria agrícola mediante el seguimiento de líneas de cultivo basado en un sistema de visión estéreo. Rovira-Más y col. (2003, 2008, 2010) han aplicado y extendido con éxito las técnicas de visión estéreo a otras áreas dentro de la AP, siempre con el propósito de obtener información 3D de la escena tanto para guiado como para

detección de elementos de interés. El procesamiento de imágenes estereoscópicas se utiliza para determinar las posiciones tridimensionales (3D) de aquellos objetos de interés obtenidos mediante el par de imágenes estereoscópicas (Guerrero y col., 2011; Pajares y col., 2012). Estas posiciones 3D, determinadas mediante el cálculo de los valores de disparidad medios del par estéreo, proporcionan la información básica para crear un mapa de elevación. El mapa resulta ser una estructura bidimensional (2D) de la misma dimensión espacial que cualquiera de las imágenes que componen el par estereoscópico y cuyos valores en cada posición indica la altura del cultivo con respecto al suelo para el píxel de imagen correspondiente a dicha posición en el mapa. Por este motivo, estos métodos requieren que el cultivo tenga una altura considerable con respecto al suelo. Debido a que en los campos de maíz, durante la fase de tratamiento, las alturas no son relevantes, esta técnica se hace ineficiente desde el punto de vista de la aplicación que se propone en el presente trabajo de investigación.

2.5.5 Métodos basados en regiones

Estos métodos encuentran y caracterizan las regiones de píxeles contiguos con el mismo valor dentro de una imagen binaria (Fontaine y Crowe, 2006). El algoritmo busca zonas etiquetadas como blancas en la imagen binaria (entre las líneas de cultivo) de más de 200 píxeles, bajo el supuesto de que zonas más pequeñas pueden ser ruido en las líneas de cultivo. Una vez se han identificado las zonas, el algoritmo determina el ángulo de sus ejes principales y localiza su centro de gravedad. Para una línea recta blanca perfecta, el centro de gravedad de la zona se sitúa sobre la línea central de la línea recta blanca, y el ángulo representativo es el ángulo de los espacios entre líneas de cultivo. El algoritmo devuelve el ángulo y el centro de gravedad de la zona más cercana al centro de la imagen. La identificación de las zonas en áreas con malas hierbas, no distingue entre las zonas de mala hierba y cultivo, por lo que no es posible aplicar esta técnica cuando el propósito es el de diferenciar las líneas de cultivo de las malas hierbas como corresponde a la investigación abordada en el presente trabajo de tesis.

2.5.6 Métodos basados en la acumulación de plantas verdes

Olsen (1995) propuso un método basado en la consideración de que en una imagen, a lo largo de una línea de cultivo aparece una importante acumulación de partes verdes. La imagen se transforma a escala de grises de manera que las partes verdes

aparecen más claras que el resto. Se buscan acumulaciones de píxeles explorando todas las columnas de una determinada zona rectangular. Esta exploración se basa en la hipótesis de que las líneas verticales siguen esta dirección en la imagen. Posteriormente se ajustan una serie de rectas, por medio de mínimos cuadrados, sobre las acumulaciones de píxeles previamente obtenidas. Los máximos locales de las acumulaciones proporcionan las posiciones de los centros de las líneas. El método propuesto se aplica sobre imágenes que carecen de proyección de perspectiva dado que son adquiridas con una cámara situada en posición ortogonal respecto al suelo.

2.5.7 Métodos basados en análisis de frecuencia

Dado que las líneas de cultivo son verticales en la escena 3D, éstas se proyectan sobre la imagen bajo la proyección de perspectiva del sistema óptico; por tanto, la regularidad espacial 3D permite suponer la existencia de un determinado comportamiento también regular en el dominio de la frecuencia. Vioix y col. (2002) aprovechan esta característica y aplican un filtro de Gabor bidimensional, definido como un ajuste de una función Gaussiana por una señal coseno. El parámetro de frecuencia necesario para el filtro de Gabor se deduce empíricamente mediante la aplicación de una transformada rápida de Fourier bidimensional (Bossu y col., 2006).

Bossu y col. (2009) aplican técnicas basadas en la transformada de *wavelets* para discriminar líneas de cultivo basándose en el posterior análisis de frecuencia para los coeficientes horizontales, verticales y diagonales que proporciona dicha transformada. Se basan en el hecho de que las líneas de cultivo están bien ubicadas espacialmente sobre la escena, de modo que mediante la transformada aparecerán patrones característicos sobre las imágenes a una frecuencia determinada; de este modo, seleccionando una función *wavelet* madre con esta frecuencia se pueden identificar las líneas de cultivo. El cultivo, en las imágenes obtenidas durante la fase experimental recogida en este trabajo de tesis, no muestra un claro contenido de frecuencia en el espacio de Fourier, por lo tanto la aplicación de otros filtros basados en la frecuencia supone una tarea difícil de llevar a cabo.

2.5.8 Métodos basados en regresión lineal

Billingsley y Schoenfisch (1997) presentan un sistema de detección de cultivo que es relativamente invariable al ruido extra provocado por la existencia de rodales de malas hierbas. Utilizaron regresión lineal en cada uno de los tres segmentos considerados de líneas de cultivo, y una función de coste similar en el momento de encajar la línea de mejor ajuste con los valores atípicos (es decir, ruido generado por las malas hierbas) como un medio de identificar la información para el guiado a través de una línea. En Montalvo y col. (2012a) se aplica una regresión lineal para la detección de líneas de cultivo en imágenes con una alta densidad de malas hierbas. La regresión lineal también se aplica en Søgaaard y Olsen (2003), pero una regresión lineal simple, tal y como se verá en el capítulo tres sección 3.5.1, es altamente sensible a la aparición de rodales de malas hierbas aislados entre las líneas de cultivo y también a rodales que estén solapados con el cultivo. Por ello, en el capítulo tres sección 3.5.2 se describe un avance de la técnica presentada en el capítulo cinco sección 5.4.3. Esta técnica de regresión lineal está basada en el estimador Theil-Sen (Theil, 1950; Sen, 1968) el cual no se ve influido por la aparición de puntos entre líneas, dando lugar a un algoritmo robusto, automático y que ha sido probado en términos estadísticos con resultados muy satisfactorios dentro de los experimentos llevados a cabo dentro del marco de este trabajo de tesis.

Procesamiento de imágenes en agricultura de precisión

3.1 Introducción

Uno de los intereses principales de los métodos de procesamiento de imágenes digitales es el procesamiento de los datos para la percepción de las máquinas de forma autónoma (Pajares y Cruz, 2007).

En este capítulo se detalla el conjunto de técnicas que se integran en los sistemas desarrollados para la identificación de líneas de cultivo y malas hierbas en campos de cultivo de maíz con el fin de realizar un tratamiento selectivo de las infestaciones así como guiar el movimiento del tractor. El esquema general propuesto para identificar las líneas de cultivo y malas hierbas que aparecen en una imagen se muestra en la figura 3.1.

Como puede apreciarse, sobre cada imagen perteneciente a un cultivo de maíz, se realiza una serie de procesamientos para extraer la capa vegetal e identificar las líneas de cultivo así como las malas hierbas. Una de las primeras fases consiste en la aplicación de técnicas de extracción de verdes y de umbralización con el fin de obtener una imagen binaria, en blanco y negro, donde el blanco se corresponde con la capa vegetal (cultivo y malas hierbas) y el negro con el suelo, cielo y el resto de elementos que aparecen en la imagen. Una vez se tiene identificada la cubierta vegetal, es necesaria la detección o identificación de las líneas de cultivo como paso previo para la posterior discriminación de las malas hierbas frente al cultivo.

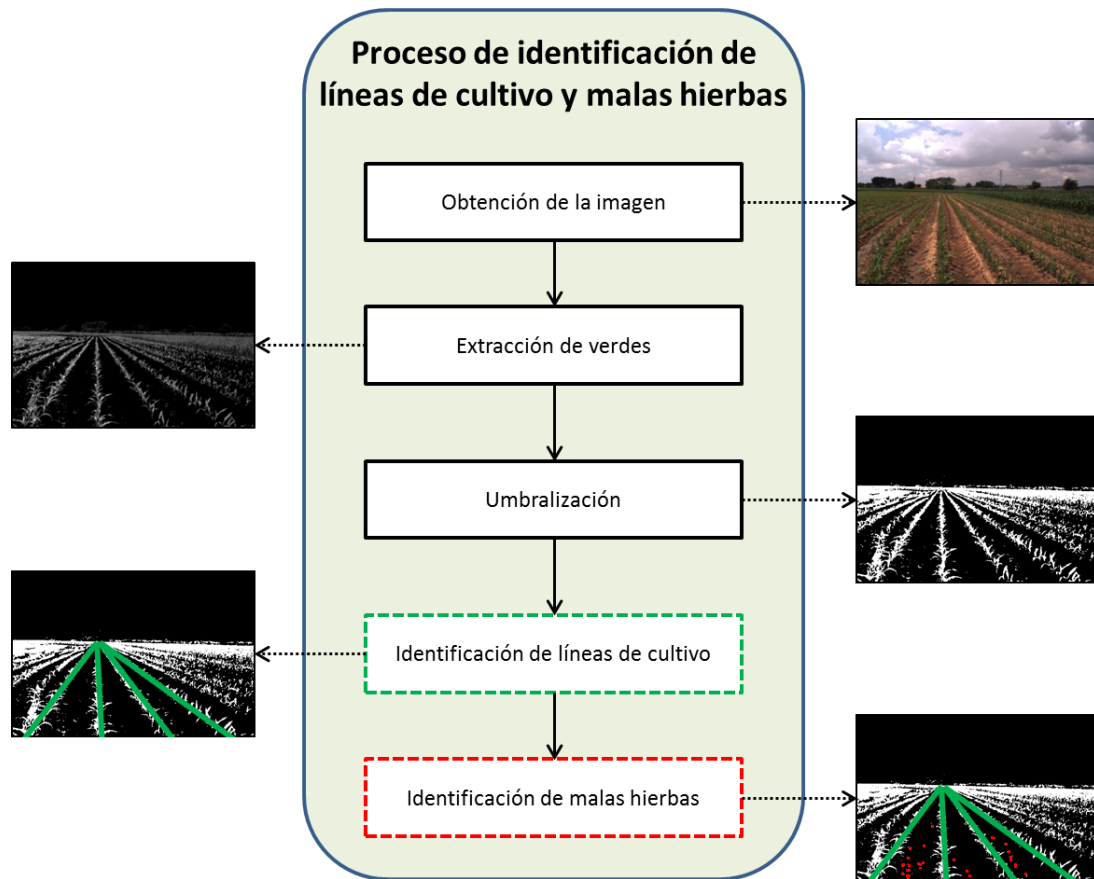


Figura 3.1 Esquema del proceso general para la identificación de líneas de cultivo y malas hierbas.

En este capítulo se analiza cada uno de los módulos que conforman el proceso mostrado en el esquema de la figura 3.1. La descripción de los módulos sirve a su vez como introducción a los capítulos cuatro y cinco donde se aportarán soluciones y nuevas técnicas con el fin de mejorar los procesos incluidos en los módulos mencionados. Así, la obtención de la imagen, la extracción de verdes y la umbralización se describen en el capítulo cuatro, mientras que los procesos de identificación de líneas y malas hierbas se explican en el capítulo cinco. En la sección 3.2 se estudian diferentes modelos de color para cámaras digitales así como el conocido efecto *vignetting* que suele aparecer al trabajar con este tipo de cámaras conjuntamente con algunos elementos anexos como son el sistema óptico y los filtros de corte de la radiación ultravioleta e infrarroja. En la sección 3.3 se describen diferentes índices orientados a identificar la capa vegetal a partir de una imagen dada, que son los utilizados en los métodos propuestos en este trabajo. En la sección 3.4 se estudian diferentes métodos de umbralización, en concreto se detalla el método de Otsu (1979), así como transformaciones morfológicas que permiten eliminar zonas irrelevantes de la imagen.

Por último, en la sección 3.5 se introduce la técnica, posteriormente descrita en detalle en el capítulo cinco, sección 5.4.3, con el fin de identificar y corregir los errores en la detección de las líneas de cultivo así como la identificación de las malas hierbas para su posterior tratamiento.

3.2 Imágenes: modelos de color y efecto *vignetting*

Una imagen es una representación de la escena que manifiesta la apariencia visual de un objeto real en dicha escena. Esta imagen puede provenir de diferentes fuentes, y tal y como se ha mencionado en el capítulo dos sección 2.2.3, las imágenes utilizadas para la investigación desarrollada provienen de cámaras digitales equipadas con sensor CCD. Como bien es sabido, las imágenes digitales son una representación bidimensional de una escena tridimensional. Dichas imágenes, en color, se obtienen mediante el procedimiento descrito en la sección 2.2.4, resultando finalmente un conjunto de tres matrices numéricas, una por cada canal de color cuyos valores difieren en función del modelo de color que se utilice en cada caso. A continuación se describen los modelos de color, junto con sus posibilidades de aplicación en el ámbito de la investigación desarrollada para imágenes de cultivos de maíz. También se estudia el efecto *vignetting*, junto con sus causas, que afecta a la composición del color por la modificación de los valores numéricos de las matrices de color.

3.2.1 Modelos de color

Un modelo de color es un modelo matemático abstracto que describe la forma en que se representan los colores como matrices de números, utilizando normalmente tres o cuatro valores o componentes de color. Cuando un modelo se asocia a una descripción precisa de cómo se han de interpretar las componentes, el conjunto de colores obtenido se denomina espacio de color. A continuación se describen distintas formas en las que la visión a color puede ser representada.

3.2.1.1 RGB

El modelo de color RGB (*Red, Green, Blue*; Rojo, Verde, Azul) está basado en la síntesis aditiva, mediante la cual los colores primarios rojo (R), verde (G) y azul (B) se combinan de distintas formas para generar un conjunto de colores. Así, combinando uno de estos colores primarios con otro también primario en proporciones iguales se

obtienen los colores aditivos secundarios: cian (C), magenta (M) y amarillo (Y). Combinando los tres colores primarios con los máximos valores de representación se obtiene el blanco puro (W), si la combinación se realiza con los mínimos valores posibles el resultado es el negro puro y si la combinación se realiza también con idénticas proporciones de colores primarios pero con valores de representación intermedios, se obtienen los grises. Este modelo de color es uno de los más utilizados en las cámaras a color. Un canal de color está representado por una matriz donde el valor de cada componente se calcula a partir de una combinación de valores digitales en bits, así en una representación de 8 bits por valor, el máximo valor de representación es $2^8 = 256$ y el mínimo es 0. En este caso, el rango de valores queda establecido entre 0 y 255, en el que se sitúan los valores intermedios. La figura 3.2 muestra el esquema general del modelo de color RGB con los colores primarios, secundarios y el blanco.

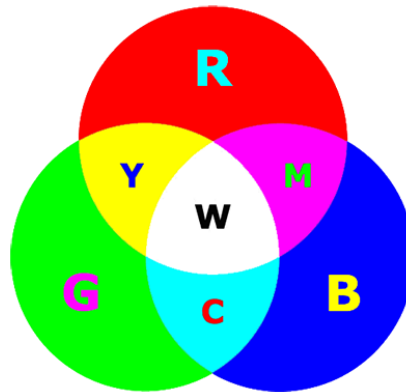


Figura 3.2 Esquema del modelo de color RGB.

3.2.1.2 HSI

En el modelo HSI (*Hue, Saturation, Intensity*; Matiz, Saturación, Intensidad) los colores se distinguen unos de otros por su tono o matiz, intensidad y saturación. El tono hace referencia al color dominante tal y como se percibe, la intensidad representa la iluminación percibida, y la saturación es la cantidad de luz blanca mezclada con el color dominante, y por lo tanto, diferencia un color intenso de uno pálido. Es por eso que este modelo se utiliza a menudo cuando la iluminación de la escena es determinante y se desea trabajar con la intensidad de la imagen. Existe una relación de transformación entre los modelos RGB y HSI y viceversa (Pajares y Cruz, 2007).

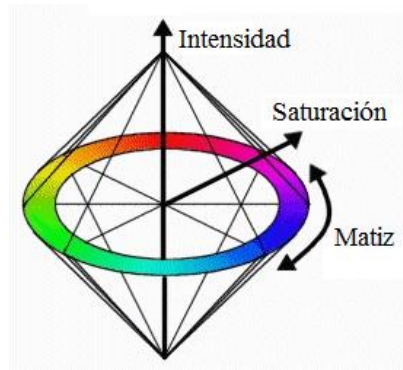


Figura 3.3 Esquema del modelo de color HSI.

3.2.1.3 YUV

El modelo YUV representa la percepción humana del color de una forma más cercana que el estándar RGB usado por el hardware de gráficos por ordenador, ya que considera la característica de la visión humana de ser más sensible a los cambios en la intensidad de la luz que a los cambios en el color. Se utiliza porque admite un nivel razonable de error sin que el humano aprecie el fallo. El parámetro Y representa la luminancia (información en blanco y negro), mientras que U y V representan la crominancia (información relacionada con el color). Este modelo se desarrolló para codificar la transmisión de información de los televisores a color garantizando simultáneamente el funcionamiento de los televisores en blanco y negro, ya que permite enviar información de color independiente de la información de luminancia. Existe una relación de transformación entre los modelos RGB y YUV y viceversa (Dawson-Howe, 2014).

3.2.1.4 $YD_B D_R$

Basado en el modelo YUV, este modelo es utilizado en el estándar de televisión en color SECAM. Como en el caso anterior consta de una componente que representa la luminancia (Y) y dos que representan la crominancia (D_B y D_R). En este caso se relaciona con el modelo RGB en que los valores se suman para generar una señal simple de luminancia Y, mientras que la señal D_B se obtiene restando Y de la señal azul del modelo RGB, y D_R restando Y de la señal roja, siendo ambas multiplicadas por factores de escala diferentes. Es por ello que este modelo es útil para diferenciar los píxeles de una imagen con dominancia de la componente roja.

3.2.2 Efecto *vignetting*

Al trabajar con cámaras digitales, un fenómeno que puede manifestarse es el llamado efecto *vignetting*, el cual se caracteriza por la aparición de bordes negros o sombreados en las esquinas de las fotografías. Este efecto puede darse de dos formas distintas (Jacobson y col., 2000):

a) *Vignetting* mecánico, provocado por el bloqueo de la luz debido a parasoles mal colocados, el uso excesivo de filtros o el uso de objetivos no acordes al tamaño del sensor de la cámara, donde el chasis o el soporte metálico del filtro o portafiltros entra en el ángulo de visión del objetivo y sale en la fotografía, figura 3.4(b). Este tipo de *vignetting* es apreciable y fácilmente corregible adaptando las lentes, filtros, parasoles, u otros dispositivos a la cámara que se utiliza para conseguir una proyección sin obstáculos, figura 3.4(a).

b) *Vignetting* óptico, a diferencia del mecánico, éste presenta un degradado continuado desde el centro de la imagen hacia el exterior. Se puede apreciar al fotografiar una superficie homogénea. Este *vignetting* puede ser provocado por la apertura del iris al máximo, por la construcción del objetivo y disposición de las lentes en relación al ángulo de incidencia de la luz, o por el uso de filtros creados por múltiples capas superpuestas, ya que los rayos perpendiculares al centro recorren un espacio menor que los que inciden cerca de los bordes, es por esto que este efecto se aprecia más cuanto mayor es la distancia al centro de la imagen, figura 3.4(c).

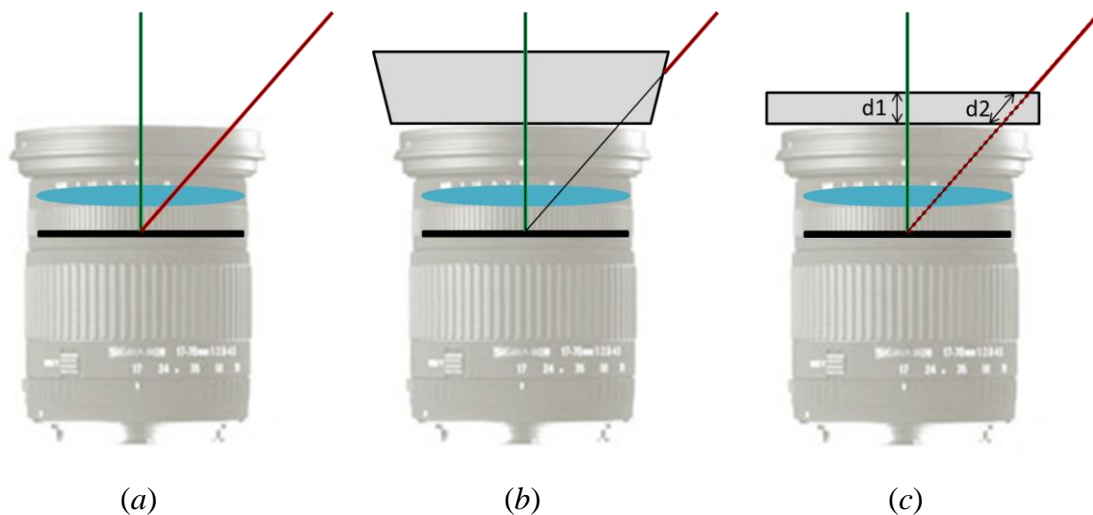


Figura 3.4 (a) Incidencia de los rayos sin interposición de filtros o parasoles sobre la lente.
 (b) *Vignetting* mecánico, incidencia de los rayos sobre una lente con un parasol mal colocado.
 (c) *Vignetting* óptico, incidencia de los rayos sobre una lente con un filtro multicapa.

Como se puede observar en la figura 3.4(b), el *vignetting* mecánico se debe a un mal uso de parasoles o filtros cuyas partes opacas se interponen dentro del ángulo de incidencia de los rayos. Este tipo de *vignetting* también aparece cuando se usa una lente cuyo campo de proyección no cubre íntegramente las dimensiones del sensor CCD, ya que si se usa una lente de menor tamaño que el sensor, ésta no es capaz de estimular todo el sensor y se aprecia una distorsión acentuada particularmente en las esquinas. Por otro lado, en la figura 3.4(c) se puede ver cómo la distancia que recorre el rayo lateral (d_2) es mayor que la recorrida por el rayo central (d_1), esto provoca que los sensores en sus partes más externas y especialmente en las esquinas estén menos estimulados que en la parte central, apareciendo en este caso el llamado *vignetting* óptico. Este efecto se utiliza en ocasiones en fotografías artísticas. En el caso del trabajo de investigación presentado es un efecto no deseado puesto que afecta mucho a la componente verde esencial en los algoritmos de identificación de capa vegetal, figura 3.5(b).

Las imágenes de las figuras 3.5 (a) y (b) muestran ejemplos de los efectos *vignetting*, mecánico y óptico respectivamente.

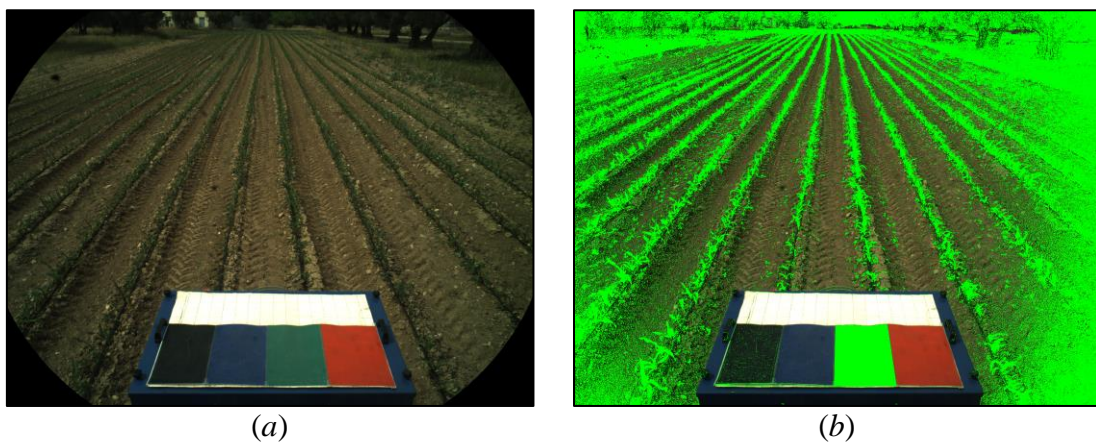


Figura 3.5 Imágenes con efecto *vignetting*: (a) mecánico, provocado por el uso de un filtro cuyo tamaño no se ajusta al sistema; (b) óptico, al aplicar un algoritmo de extracción de la capa vegetal.

Como se aprecia, en el caso de un *vignetting* mecánico el efecto se produce sólo en las esquinas y no repercute sobre el resto de la imagen, mientras que en el caso de un *vignetting* óptico, éste afecta a toda la imagen, ya que el efecto es gradual propagándose desde el centro de la imagen hacia los laterales y especialmente en las esquinas, que es donde más se acentúa, debido a la mayor distancia recorrida por los rayos incidentes sobre el filtro. La imagen de la figura 3.5(a) ha sido tomada con la cámara SVS-VISTEK (2014) modelo SVS4050CFLGEA en la que se ha instalado un filtro que bloquea parte del sensor, mientras que la imagen de la figura 3.5(b) ha sido tomada con

la misma cámara pero instalando el filtro de corte Schneider UV/IR 486 (Schneider, 2014b). Se puede apreciar cómo al aplicar un algoritmo de extracción de la capa vegetal aparece un degradado continuo hacia los laterales de la imagen que se acentúa en las esquinas. Los píxeles marcados en verde representan los píxeles de la imagen original identificados como pertenecientes a plantas verdes, por lo que claramente origina falsos positivos que han de corregirse o eliminarse con el fin de minimizar los errores posteriores en la detección. Este filtro se fabrica mediante la superposición de diferentes capas de material, ya que se trata de un filtro de corte del IR, en las partes alejadas del centro de la imagen se produce un mayor filtrado del IR afectando también en parte a la componente roja. Esto hace que la componente verde predomine sobre la roja que ha sido parcialmente atenuada por el filtro. El resultado es el exceso de verde que aparece en las imágenes afectadas por el *vignetting*. Otra característica de este filtro es que bloquea la incidencia sobre el sensor de luz ultravioleta (UV) a 300-375nm e infrarroja (IR) a 750-1035nm, por lo que permite la entrada de radiaciones con longitudes de onda en el rango $390\text{nm} \pm 15\text{nm}$ a $690\text{nm} \pm 20\text{nm}$, tal y como se especifica en la curva de respuesta característica del filtro. Las imágenes agrícolas, objeto de la investigación realizada, se caracterizan porque poseen zonas de suelo con plantas verdes (cultivo y malas hierbas). Los suelos agrícolas, suelen ser arcillosos con predominio de la componente roja en su tonalidad de color. Esto hace que, si incide luz infrarroja sobre el sensor, la componente roja (R) se vea incrementada en detrimento de las otras dos componentes (G y B) apareciendo una imagen con tendencia excesiva hacia el rojo lo que enmascara la vegetación, figura 3.6(a). Para corregir esta situación es necesaria la utilización de un filtro de corte del tipo indicado anteriormente, figura 3.6(b).



(a)



(b)

Figura 3.6 (a) Imagen de líneas de cultivo tomada con la cámara SVS-VISTEK sin filtro de corte UV/IR; (b) la misma imagen con filtro de bloqueo UV/IR.

Si bien, el hecho de incorporar tal filtro provoca el mencionado efecto *vignetting*, el cual requiere de una corrección adecuada para eliminar o suavizar el degradado de color que se produce. Esta corrección se realiza mediante un procesamiento de la imagen, que se detalla en el capítulo cuatro sección 4.2.2.1.

3.3 Extracción del verde

Un punto clave en AP es la identificación de las plantas, ya que la mayoría de las técnicas requieren de esta tarea para diferentes propósitos. Es por ello que el éxito de los tratamientos específicos en AP depende directamente de la correcta identificación de dichas plantas, que en el caso de la investigación que nos ocupa, pertenecen al cultivo y a las malas hierbas. Asimismo, en post-emergencia y en época de tratamiento tanto el cultivo como la mala hierba aparecen en el campo con un color predominantemente verde, aunque bien es cierto que en ocasiones con diversas tonalidades.

Diversas técnicas utilizadas en AP hacen uso de índices que tienen en cuenta el canal infrarrojo con el fin de medir el índice de biomasa de una zona por la extracción de los verdes existentes en la misma. Estos índices permiten estimar y evaluar el estado de salud de la vegetación en base a la medición de la radiación que las plantas emiten o reflejan. Un ejemplo de este tipo de índices es el NDVI (*Normalized Difference Vegetation Index*; índice de vegetación diferencial normalizado) propuesto por Rouse y col. (1973) y que se calcula mediante la ecuación (3.1). El índice NDVI se utiliza sobre todo para mejorar la discriminación de las capas vegetales, medir el vigor de las plantas y la producción de biomasa. El estudio del comportamiento radiométrico de la vegetación, que está íntimamente relacionado con la actividad fotosintética y la estructura foliar de la planta, permite determinar si la vegetación se encuentra sana o si por el contrario manifiesta algún grado de estrés.

$$NDVI = \frac{IRC - R}{IRC + R} \quad (3.1)$$

donde *IRC* es la reflectividad del infrarrojo cercano y *R* es la reflectividad en el canal rojo.

En general, una capa vegetal en buen estado de salud se caracteriza por el contraste entre la banda del rojo (entre 0.6 y 0.7 μm), la cual es absorbida en gran parte

por las hojas, y el infrarrojo cercano (entre 0.7 y 1.1 μm), que es reflectada en su mayoría debido a que las paredes de las células de las hojas poseen un alto porcentaje de agua. Cuando la vegetación sufre algún tipo de estrés, ya sea por presencia de plagas o por sequía, la cantidad de agua disminuye en las paredes celulares por lo que la reflectividad disminuye el valor IRC y aumenta paralelamente en el rojo al tener menor absorción clorofílica. El rango de valores de las reflexiones espectrales se encuentra entre el 0 y el 1, ya que, tanto la reflectividad del infrarrojo cercano como la del rojo, son cocientes de la radiación reflejada sobre la radiación entrante en cada banda espectral. Como consecuencia de estos rangos de valores, el NDVI varía entre -1 y 1.

Además del NDVI existen otros índices para estimar la capa vegetal en una imagen y que utilizan el espectro infrarrojo basándose en el hecho de que la vegetación absorbe bien la radiación correspondiente al espectro rojo visible y refleja eficientemente el infrarrojo cercano. Las ecuaciones de (3.2) a (3.5) definen cuatro de estos índices.

<i>Environmental Vegetation Index</i> (Birth y McVey, 1986)	$EVI = \frac{IRC}{R}$	(3.2)
<i>Enhanced Vegetation Index (MODIS)</i> (Huete y col., 2002)	$EVI = g \cdot \frac{IRC - R}{IRC + C_1 \cdot R - C_2 \cdot B + L}$ con $L=1$, $C_1 = 6$, $C_2 = 7.5$, y $g = 2.5$	(3.3)
<i>Perpendicular Vegetation Index</i> (Richardson y Wiegand, 1977)	$PVI = \sqrt{(R_{suelo} - R_{veg})^2 + (IRC_{suelo} - IRC_{veg})^2}$	(3.4)
<i>Soil-Adjusted Vegetation Index</i> (Huete, 1988)	$SAVI = (1 + L) \frac{(IRC - R)}{IRC + R + L}$	(3.5)

En el índice EVI (*Enhanced Vegetation Index*, MODIS), el parámetro L realiza un ajuste de la capa vegetal que relaciona de forma no lineal el infrarrojo cercano con el espectro rojo radiante que se transfiere a través de la capa vegetal. C_1 y C_2 son coeficientes que se utilizan para corregir la influencia del espectro azul en el espectro rojo y g es el factor de ganancia (Huete y col., 1994, 1997). Por el contrario, en SAVI (*Soil-Adjusted Vegetation Index*), L es un factor de corrección, que varía desde 0 (alta cobertura vegetal) hasta 1 (baja cobertura) y se utiliza para corregir la alta reflectancia del suelo en regiones áridas.

Resulta de particular relevancia el NDVI, que posee un gran valor en términos ecológicos, ya que es un buen estimador de la fracción de la radiación fotosintéticamente activa interceptada por la vegetación (fPAR) (Monteith, 1981), la productividad primaria (Paruelo y col., 1997; Tucker y col., 1985) y una variable integradora del funcionamiento del ecosistema (Virginia y Wall, 2001). Posee la ventaja de tener una gran sencillez de cálculo y de facilitar la interpretación directa de los parámetros biofísicos de la vegetación. Si bien, tiene el inconveniente de que posee poca capacidad de minimizar la influencia del suelo y de la atmósfera. Es en relación a este aspecto donde se hace ineficiente en el problema que se aborda en el presente trabajo de investigación, ya que resulta absolutamente necesario el hecho de poder diferenciar claramente las plantas del suelo y otros elementos con el fin de identificar el cultivo y las malas hierbas con la máxima precisión posible. Además, se hace necesario el uso de cámaras multiespectrales en lugar de cámaras RGB, lo que encarece los sistemas suponiendo una desventaja de cara a la posible comercialización del sistema desarrollado. Algunas propuestas recientes permiten la obtención simultánea de los valores espectrales necesarios para calcular del índice NDVI a partir de la adecuada manipulación de los filtros de cámaras comerciales como la Canon 500D y la Canon 350D (Rabatel y col., 2014)

En definitiva de lo anterior, se desprende que resultaría muy interesante disponer de índices que únicamente trabajen con el espectro visible, con el fin de obtener una imagen en escala de grises donde las plantas y el suelo aparezcan muy contrastados. Por ejemplo, el índice diferencial normalizado NDI (*Normalized Difference Index*), propuesto en Perez y col. (2000), pese a estar basado en el NDVI, únicamente utiliza los canales verde (G) y rojo (R) haciendo posible trabajar con una imagen en el modelo RGB.

$$NDI = \frac{G - R}{G + R} \quad (3.6)$$

Es por ello que a lo largo de este trabajo se analizan y utilizan una serie de índices de vegetación con el fin de extraer los píxeles de una imagen que pertenecen a la capa vegetal. Como puede verse en Meyer y Camargo-Neto (2008), la ventaja de utilizar índices de color es que éstos acentúan un color en particular, el cual es de interés para el estudio, en este caso, el color verde. Las imágenes que se utilizan contienen tres zonas dominantes principales, las cuales pueden identificarse con cada canal de la

imagen RGB: el verde para las plantas, rojo para el suelo, y en ocasiones azul para el cielo. Desde este punto de vista, los índices de verde para identificar las plantas y los índices de rojo para la identificación del suelo, resultan apropiados.

Para el uso de índices basados en componentes del espectro visible es necesario realizar una normalización de los valores de cada canal. Por ello, dada una imagen original en el espacio de color RGB, se aplica la relación de normalización, ecuación (3.7), utilizada habitualmente en la segmentación de imágenes agrícolas (Gée y col., 2008), obteniendo las componentes espectrales r , g y b normalizadas en el intervalo $[0, 1]$.

$$r = \frac{R_n}{R_n + G_n + B_n}, \quad g = \frac{G_n}{R_n + G_n + B_n}, \quad b = \frac{B_n}{R_n + G_n + B_n} \quad (3.7)$$

donde R , G y B son las componentes RGB normalizadas en el rango de 0 a 1, obtenidas de la siguiente forma:

$$R_n = \frac{R}{R_{\max}}, \quad G_n = \frac{G}{G_{\max}}, \quad B_n = \frac{B}{B_{\max}} \quad (3.8)$$

donde $R_{\max} = G_{\max} = B_{\max} = 255$ para imágenes a color en una representación de 24-bit, que se corresponde con 8 bits por canal y por píxel.

Los índices de vegetación se pueden utilizar independientemente o combinados, tal como se propone en Guijarro y col. (2011). Las ecuaciones de (3.9) a (3.12) corresponden a cuatro índices muy utilizados en el tratamiento de imágenes agrícolas.

<i>Excess Green</i> (Woebbecke, 1995; Ribeiro y col., 2005)	$ExG = 2g - r - b$	(3.9)
<i>Color index of vegetation extraction</i> (Kataoka y col., 2003)	$CIVE = 0.441r - 0.811g + 0.385b + 18.78745$	(3.10)
<i>Vegetative</i> (Hague y col., 2006)	$VEG = \frac{g}{r^a b^{1-a}}, \text{ con } a = 0.667$	(3.11)
<i>Excess green minus excess red</i> (Neto 2004; Meyer y Camargo-Neto 2008)	$ExGR = ExG - ExR$	(3.12)

donde el exceso de rojo se calcula mediante la siguiente expresión (Meyer y col. 1998): $ExR = 1.4r - g$. En Guijarro y col. (2011) los cuatro índices anteriores se combinan para obtener el índice COM tal como se muestra en la expresión (3.13).

$$COM = w_{ExG}ExG + w_{ExGR}ExGR + w_{CIVE}CIVE + w_{VEG}VEG \quad (3.13)$$

donde $w_{ExG} = 0.25$, $w_{ExGR} = 0.30$, $w_{CIVE} = 0.33$ y $w_{VEG} = 0.12$ son los pesos, obtenidos experimentalmente para cada índice, y que representan la importancia relativa del índice en la combinación. La imagen resultante de la combinación, COM, se transforma linealmente al rango de valores en el intervalo $[0, 1]$.

En la figura 3.7, se muestra un ejemplo de extracción de verdes mediante el índice COM a partir de una imagen en RGB, figura 3.7(a), dando lugar a la imagen en escala de grises de la figura 3.7(b).



Figura 3.7 (a) Imagen original. (b) Extracción de verdes mediante el índice COM (Guijarro y col., 2011).

3.4 Umbralización

Los índices estudiados en el apartado anterior transforman las imágenes de color a escala de grises, es por ello que se necesita determinar un umbral a partir del cual un píxel es considerado planta u otro material para conseguir una imagen binaria donde los píxeles pertenecientes a la vegetación se etiqueten con blancos y el resto en negro.

La mayoría de estas técnicas utilizan el histograma de la imagen, figura 3.8(b), obtenido a partir de una imagen en escala de grises, figura 3.8(a). El histograma es una representación gráfica de la distribución de los distintos valores de intensidad de una imagen. El número de veces que un determinado valor aparece en la imagen se

denomina frecuencia de dicho valor. En una representación gráfica en forma de barras, la altura de cada barra es proporcional a la frecuencia de los valores representados, es decir, cada barra representa la frecuencia de aparición de cada valor de intensidad. Una vez obtenido el histograma, pueden aplicarse diferentes técnicas con el fin de obtener un valor de umbral adecuado (Gonzales-Barron y Butler, 2006; Sahoo y col., 1988).

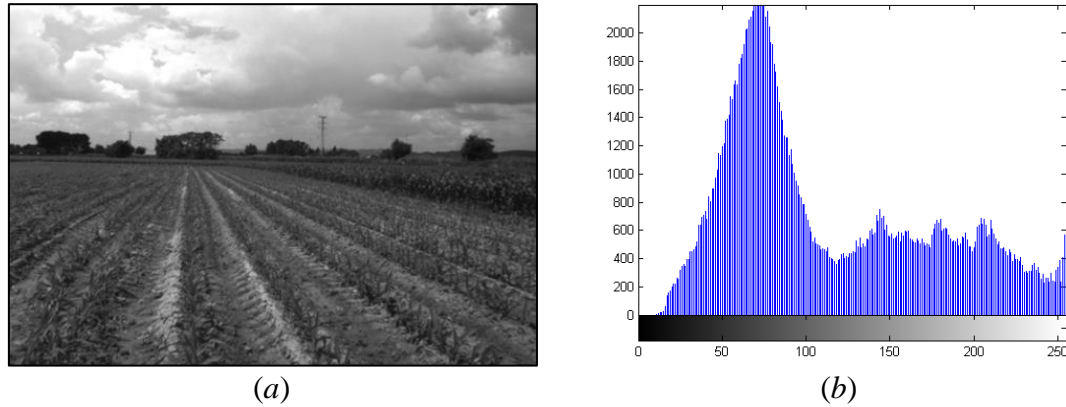


Figura 3.8 (a) Imagen en escala de grises obtenida a partir de la figura 3.7(a). (b) Histograma obtenido.

3.4.1 Umbralización manual

Conocido el histograma se puede realizar una umbralización manual a partir de un valor fijo, o bien proceder tal y como se hace en Gonzales-Barron y Butler (2006), donde la forma de seleccionar un umbral manual consiste en identificar los picos del histograma y los valles entre ellos. Una vez localizados ambos, el umbral adecuado es aquel que se corresponde con el valor de frecuencia que resulta ser un mínimo local y está situado inmediatamente a la izquierda del valor medio de intensidad.

En ambos casos, al igual que ocurre en el método propuesto en Romero y col. (2012b), se requiere una intervención por parte del usuario donde se analice cada imagen para seleccionar el umbral correcto, por lo que el uso de esta técnica hace inviable su aplicación en tareas automáticas en tiempo real y en entornos de exterior donde las condiciones de iluminación son muy variables. Es por ello que es necesario aplicar técnicas de umbralización automática entre las que se pueden encontrar: algoritmo de mínimo error (Kittler y Illingworth, 1986), umbralización *fuzzy* (Huang y Wang, 1995), algoritmo *K-means* (Hartigan, 1975), método de entropía (Pun, 1980), algoritmo de preservación del momento (Tsai, 1985), ISODATA (Ridler y Calvard, 1978) o el método de Otsu (1979). A continuación se describe este último por ser el finalmente elegido, dado su rendimiento y eficacia en las imágenes utilizadas en este trabajo de investigación.

3.4.2 Método de Otsu

Uno de los métodos de umbralización más conocido es el método de Otsu (1979), aplicado en diferentes tareas (Nixon y Aguado, 2002) incluyendo tareas agrícolas como las que aquí se tratan (Guerrero y col., 2012a, b, 2013; Montalvo y col., 2012a, b, 2013a; Romeo y col., 2013a; Burgos-Artizzu y col., 2011; Guijarro y col., 2011, 2013). El método de Otsu es estudiado en Meyer y Camargo-Neto (2008) demostrando ser eficaz. Además, Sezgin y Sankur (2004) ponen también a prueba la eficacia del método en imágenes donde el número de píxeles situados a ambas partes del histograma de la imagen, dividido por el umbral de Otsu, llegan a situarse muy próximos los unos de los otros.

A partir de una imagen en escala de grises el método de Otsu determina de forma completamente automática el mejor umbral para la obtención de dos clases diferenciadas, es decir, busca aquel umbral que minimiza la varianza intra-clase y maximiza la varianza inter-clase con el fin de obtener la mayor distancia media entre los centroides de dichas clases. Para ello, dada una imagen con niveles en el rango $[0, L-1]$, se definen como C_0 los píxeles con los niveles entre $[0, \dots, k]$ y C_1 aquellos con niveles en $[k+1, \dots, L-1]$, siendo k el umbral y L el nivel máximo de intensidad (256 en imágenes con representación de 8 bits por píxel, como es nuestro caso). El número de píxeles en el nivel i se define como $h(i)$ y N es el número total de píxeles de la imagen. La probabilidad de pertenencia a cada nivel de gris p_i se define como:

$$p_i = \frac{h(i)}{N} \quad (3.14)$$

Las probabilidades de pertenencia (ω) y los niveles medios (μ) de las clases en el nivel k se obtienen según las ecuaciones (3.15) y (3.16).

$$\begin{aligned} \omega_0(k) &= \Pr(C_0) = \sum_{i=0}^k p_i \\ \omega_1(k) &= \Pr(C_1) = \sum_{i=k+1}^{L-1} p_i = 1 - \omega_0(k) \end{aligned} \quad (3.15)$$

$$\begin{aligned}\mu_0(k) &= \sum_{i=0}^k i \Pr(i | C_0) = \sum_{i=0}^k i p_i / \omega_0(k) \\ \mu_1(k) &= \sum_{i=k+1}^{L-1} i \Pr(i | C_1) = \sum_{i=k+1}^{L-1} i p_i / \omega_1(k)\end{aligned}\quad (3.16)$$

Una vez se obtienen las probabilidades de pertenencia y los niveles medios de cada clase, se calcula la varianza inter-clase (σ_B^2) según la ecuación (3.17).

$$\sigma_B^2(k) = \omega_0(k) \cdot \omega_1(k) \cdot (\mu_1(k) - \mu_0(k))^2 \quad (3.17)$$

siendo el umbral óptimo k^* aquel que maximiza $\sigma_B^2(k)$, por lo tanto:

$$\sigma_B^2(k^*) = \max_{0 \leq k < L} [\sigma_B^2(k)] \quad (3.18)$$

Con ello se obtiene un umbral k^* perteneciente al intervalo $[0, 255]$. Posteriormente, la binarización se realiza comparando píxel a píxel con dicho umbral k^* , de manera que aquellos valores de la imagen en escala de grises obtenida a partir de la extracción del verde que superen el valor del umbral de Otsu pertenecen a las plantas (clase C_1), y aquellos que no lo consiguen se identifican como parte del suelo u otros sedimentos (clase C_0). Hay que tener en cuenta que la mayoría de los métodos de extracción del verde obtienen valores en el intervalo $[0, 1]$, por ello es importante que tanto el método de extracción como de umbralización trabajen con valores en el mismo rango, por lo que en ocasiones es necesario realizar una transformación de escala.

La figura 3.9 ilustra el resultado de aplicar el método de Otsu a la imagen mostrada en la figura 3.7(b).



Figura 3.9 Binarización mediante el método de Otsu de la figura 3.7(b).

3.4.3 Operaciones morfológicas

Las operaciones morfológicas son técnicas muy utilizadas en el procesamiento de imágenes. Partiendo de una imagen binaria donde un valor lógico de 1 (color blanco) pertenece al objeto de interés (las plantas en este caso) y un valor lógico 0 (color negro) pertenece al fondo (suelo y otros elementos), las operaciones morfológicas pueden mejorar los resultados obtenidos mediante el proceso de binarización directa, preservando a la vez características esenciales y eliminando aspectos irrelevantes. Es por esto que las operaciones morfológicas pueden usarse con los siguientes objetivos (Pajares y Cruz, 2007):

- a) Pre-procesamiento de imágenes (suprimir ruido, simplificar formas).
- b) Destacar la estructura de objetos (transformaciones homotópicas, extraer esqueleto, marcado de objetos, envoltura convexa, granulación, rellenado de regiones, ampliación y reducción).
- c) Mejorar el resultado de la descripción cualitativa de objetos (área, perímetro, diámetro, etc.).

Habitualmente las operaciones morfológicas se usan sobre imágenes binarias, aunque también existen operaciones morfológicas para imágenes en escala de grises.

El objetivo de aplicar estas técnicas, en el marco de la investigación desarrollada, es la eliminación del ruido provocado por los algoritmos de segmentación utilizados en la sección 3.3 así como el rellenado de regiones. De esta forma, la imagen de partida necesaria para la identificación de las líneas de cultivo y malas hierbas aparece mucho más limpia eliminando los píxeles que aparecen aislados.

El conjunto más simple de operaciones morfológicas (Serra, 1982) está formado por *erosión*, *dilatación*, *apertura* y *cierre*. Estas transformaciones morfológicas utilizan elementos estructurales (M) que se construyen alrededor de un origen local (punto representativo), siendo los más comunes aquellos conjuntos que están 4-conectados (M_4) y 8-conectados (M_8), que tienen la apariencia que se muestra en (3.19).

$$M_4 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & \bullet 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad M_8 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \bullet 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (3.19)$$

Estos elementos estructurales M , tal y como se verá a continuación, se aplican sobre cada píxel de la imagen binaria I dando como resultado una posible modificación del valor del píxel en función de la operación morfológica elegida.

3.4.3.1 Erosión

La erosión, que es una transformación no invertible (la imagen original no se recupera), se lleva a cabo según la ecuación (3.20):

$$I \otimes M = \{p \in E^2 : p + m \in I \mid \forall m \in M\} \quad (3.20)$$

siendo I la imagen binaria, M el elemento estructural utilizado y E^2 el espacio Euclídeo 2D. Por lo tanto, cada píxel p perteneciente a la imagen I se evalúa con el elemento estructural, el resultado de la erosión viene dado por los píxeles p para los cuales todos los posibles $p + m$ están en la imagen I . Esta operación tiene una serie de propiedades que pueden consultarse en Pajares y Cruz (2007).

De forma gráfica, se recorre la imagen siguiendo una dirección (normalmente de izquierda a derecha y de arriba abajo), y allí donde aparece un valor 1 lógico en el origen del elemento estructural se procede a realizar la operación morfológica, de tal forma que, si todos los unos del elemento estructural coinciden con unos en la imagen binaria, ese píxel se marca con valor 1, en caso contrario se le asigna un valor lógico 0. Esta forma de cálculo es la utilizada por Low (1991).

Una de las aplicaciones más comunes de la erosión es la eliminación de detalles irrelevantes (desde el punto de vista del tamaño), con ella se consigue eliminar pequeñas islas o píxeles aislados a costa de reducir el tamaño de los objetos.

La figura 3.10 muestra el resultado de aplicar la erosión con un elemento estructural M_8 a la imagen representada en la figura 3.9.



Figura 3.10 Erosión de la figura 3.9 utilizando el elemento estructural M_8 .

Como puede verse, la mayoría del verde identificado en la imagen de la figura 3.9 ha sido eliminado, dado que las líneas no son muy anchas y existen demasiados espacios entre las plantas de maíz.

3.4.3.2 Dilatación

Esta operación es la dual de la erosión. La dilatación es el resultado de considerar el conjunto de píxeles de todas las posibles adiciones vectoriales de pares de elementos, uno de cada conjunto I y M . Desde el punto de vista visual, la dilatación añade al objeto aquellos puntos del fondo que lindan con los bordes del mismo, y tal y como ocurre con la erosión, se trata de una transformación no invertible. La dilatación se lleva a cabo según la ecuación (3.21):

$$I \oplus M = \{p \in E^2 : p = x + m \mid \forall x \in I \mid \forall m \in M\} \quad (3.21)$$

siendo I la imagen binaria, M el elemento estructural utilizado y E^2 el espacio Euclídeo 2D.

La dilatación tiene varias propiedades que permiten su fácil implementación vía software o hardware, tales como la asociativa, invariante a la traslación y ser una transformación creciente (Pajares y Cruz, 2007; Serra, 1982; Haralick y col., 1987).

La forma gráfica de aplicar esta operación, utilizada por Low (1991), es la siguiente: se recorre la imagen siguiendo una dirección (normalmente de izquierda a derecha y de arriba abajo) y donde aparece un 1 lógico en el origen del elemento estructural, éste se evalúa, de tal modo que se ponen a 1 aquellos píxeles de la imagen binaria que coinciden con un uno en dicha posición del elemento estructural.

El resultado de aplicar la dilatación con un elemento estructural M_8 a la imagen representada en la figura 3.9 se muestra en la figura 3.11.



Figura 3.11 Dilatación de la figura 3.9 utilizando el elemento estructural M_8 .

Tal y como se ha comentado anteriormente, aquellos píxeles del suelo que lindaban con las hojas de las plantas de maíz se identifican como nuevas plantas, y por lo tanto, se aprecia un aumento de la capa vegetal al aumentar el grosor de las líneas de cultivo y el tamaño de las zonas de malas hierbas.

3.4.3.3 Apertura y cierre

La combinación de ambas operaciones (erosión y dilatación), da lugar a dos nuevas transformaciones dependiendo del orden de aplicación. Dado que la erosión y la dilatación son transformaciones no invertibles, la imagen resultado es una imagen más simplificada y menos detallada que la imagen original.

La apertura consiste en aplicar una erosión seguida de una dilatación. La apertura de una imagen I por un elemento estructural M se denota como $I \circ M$ y se define según la expresión (3.22).

$$I \circ M = (I \otimes M) \oplus M \quad (3.22)$$

El cierre, por el contrario, consiste en aplicar una dilatación seguida de una erosión. El cierre de una imagen I por un elemento estructural M se denota como $I \bullet M$ y se define según la expresión (3.23).

$$I \bullet M = (I \oplus M) \otimes M \quad (3.23)$$

El resultado de aplicar ambas transformaciones con un elemento estructural M_8 sobre la imagen representada en la figura 3.9 se muestran en la figura 3.12 (a) y (b) respectivamente.

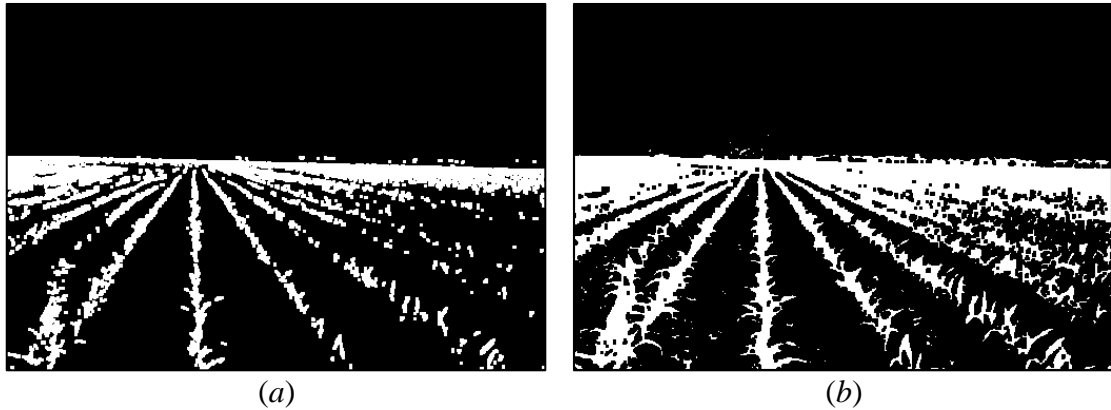


Figura 3.12 (a) Apertura. (b) Cierre.

Como se observa, estas transformaciones obtienen resultados satisfactorios en relación a la extracción de las líneas de cultivo presentes en las imágenes, siendo particularmente interesante la apertura, dado que deja únicamente aquellos conjuntos de píxeles donde existe una mayor densidad de píxeles blancos, identificándose en la realidad con aquellas zonas donde la densidad de verde es mayor, es decir, aquellos píxeles que pertenecen a líneas de cultivo o a grandes rodales de malas hierbas. Por esta razón, aplicar una apertura como paso previo a la identificación de líneas de cultivo es una técnica que facilita dicha tarea, ya que como se verá en la siguiente sección, la mayoría de los píxeles pertenecerán a una línea de cultivo y por consiguiente se puede identificar una línea de forma más rápida y eficiente a la vez que se eliminan aquellos píxeles aislados que han sido erróneamente clasificados por el método de extracción de verdes y que no resultan de interés.

3.5 Identificación de líneas de cultivo

Una vez se ha obtenido una imagen binaria, se puede estimar teóricamente la ubicación de las líneas de cultivo en la imagen mediante geometría considerando tanto los parámetros intrínsecos como extrínsecos del sistema de visión, así como la disposición de las líneas de cultivo en el campo, tal y como se verá en el capítulo cinco sección 5.4. Ahora bien, debido a que se trabaja en entornos de exterior bajo condiciones reales, las continuas vibraciones y movimientos del tractor impiden que la coincidencia entre las líneas de cultivo reales y las líneas teóricas sea máxima. En

consecuencia, con el fin de asegurar una correcta identificación de las líneas de cultivo, es necesario aplicar técnicas de ajuste basadas en los píxeles para corregir las desviaciones existentes entre las rectas estimadas y las reales. Son varios los estimadores estadísticos que se pueden utilizar para aproximar una línea recta a un conjunto de píxeles.

Así pues, el procedimiento de identificación de líneas de cultivo a partir de una imagen binaria se sintetiza como sigue:

- a) Mediante geometría, considerando los parámetros extrínsecos e intrínsecos del sistema de visión, se proyectan en la imagen las líneas rectas donde idealmente deberían situarse las líneas de cultivo reales.
- b) Por cada línea recta obtenida se realiza una búsqueda horizontal para determinar aquellos píxeles que pertenecen a la línea de cultivo dentro de un determinado ancho de búsqueda.
- c) Para cada conjunto de píxeles y dado que las líneas de cultivo se disponen en hileras, la nube de puntos tiene una tendencia lineal; por lo tanto es posible aplicar técnicas de regresión con el fin de aproximar una recta al conjunto de píxeles dado, cuyo objetivo final es la identificación correcta de las líneas de cultivo.

3.5.1 Regresión lineal simple

Una de las técnicas de regresión utilizadas es la regresión lineal simple. El análisis de regresión lineal es una técnica estadística utilizada para estudiar la relación entre una variable dependiente (Y) y las variables independientes (X).

La ecuación de una línea recta con pendiente m y término independiente b viene dada por la ecuación (3.24).

$$Y = m \cdot X + b \quad (3.24)$$

Si x_i e y_i son las coordenadas de un determinado píxel i , y N es el número de píxeles que forman parte del conjunto sobre el que se pretende ajustar la línea recta, el valor de la pendiente m y el término independiente b de la recta de regresión que mejor aproxima la nube de puntos vienen dados por las ecuaciones (3.25) y (3.26).

$$m = \frac{\left(N \cdot \sum_{i=1}^N (x_i \cdot y_i) \right) - \left(\sum_{i=1}^N x_i \cdot \sum_{i=1}^N y_i \right)}{\left(N \cdot \sum_{i=1}^N x_i^2 \right) - \left(\sum_{i=1}^N x_i \right)^2} \quad (3.25)$$

$$b = \frac{\left(\sum_{i=1}^N y_i \right) - \left(m \cdot \sum_{i=1}^N x_i \right)}{N} \quad (3.26)$$

El coeficiente de correlación lineal permite determinar si, efectivamente, existe relación lineal entre las dos variables. Una vez que se determina la existencia de relación lineal, analizando la representación gráfica de los pares de valores de ambas variables, la regresión permite definir la recta que mejor se ajusta a la nube de puntos tal y como se aprecia en la figura 3.13.

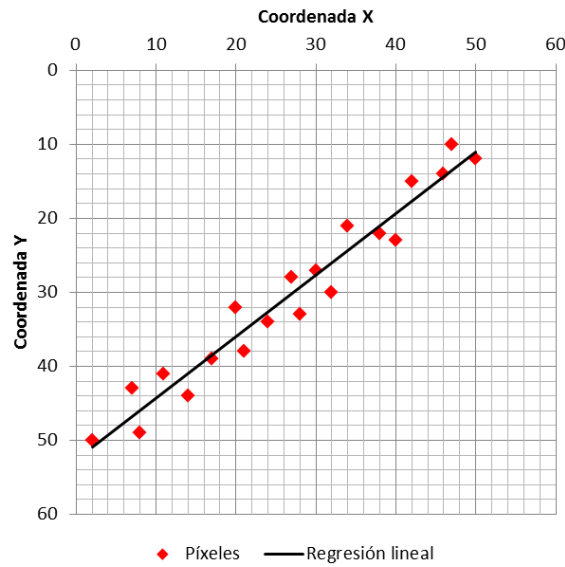


Figura 3.13 Ejemplo de recta obtenida por regresión lineal simple a partir de un conjunto de píxeles.

Respecto a los métodos de regresión lineal, tras diversos estudios en el ámbito de la investigación desarrollada, se ha determinado que en el estimador de mínimos cuadrados el coeficiente m se ve afectado por errores considerables, y además el intervalo de confianza es relativamente sensible cuando las distribuciones no son normales. Con otras medidas, como por ejemplo el punto de ruptura (Rousseeuw y Leroy, 1987), una pequeña cantidad de datos fuera de los límites pueden provocar que la estimación diverja arbitrariamente lejos de la estimación verdadera. En la aproximación propuesta

para identificar las líneas de cultivo, esto significa que unos pocos píxeles de mala hierba pueden desviar el ajuste de mínimos cuadrados lejos del ajuste de la línea de cultivo real tal y como se muestra en la figura 3.14.

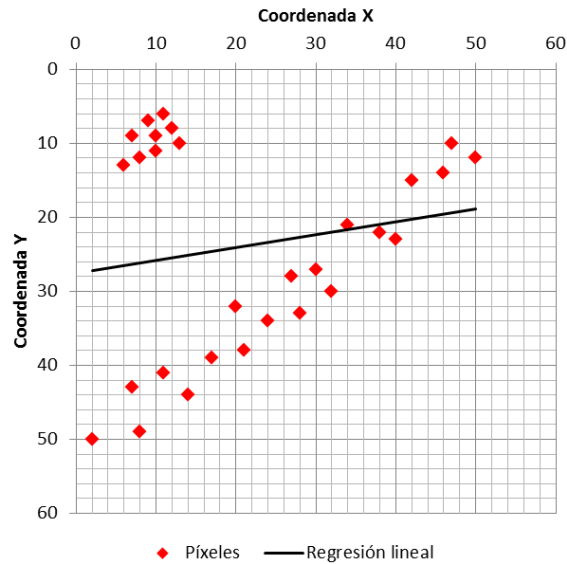


Figura 3.14 Ejemplo de recta obtenida por regresión lineal simple a partir de un conjunto de píxeles donde un grupo de píxeles de mala hierba forman parte del estudio.

Una segunda medida de la robustez la proporciona la función de influencia (Hampel y col., 1986; Huber, 1981), en la cual, el cambio en una estimación debido a la inserción de datos fuera de los límites usando una función de distancia, también puede provocar falsas estimaciones ya que la función debería tender a cero con el incremento de la distancia para ser más robusta.

Ante estos problemas expuestos, se hace necesario el uso de otros estimadores que garanticen una identificación correcta de la línea de cultivo real, evitando la desviación provocada por datos cercanos a la nube de píxeles y que no forman parte de la línea de cultivo, bien porque se trate de malas hierbas cercanas a ella o bien porque se trata de píxeles incorrectamente identificados como capa vegetal. A continuación se describe el estimador Theil-Sen que resuelve los problemas descritos.

3.5.2 Estimador Theil-Sen

Como se ha visto anteriormente, el método de identificación de líneas de cultivo debe hacer frente a las alineaciones específicas de píxeles, pero también debe ser lo suficientemente robusto como para evitar desviaciones significativas causadas por

malas hierbas que no están alineadas y se sitúan más o menos próximas a la línea de cultivo principal que se está tratando de identificar.

Otros estimadores alternativos para el cálculo del coeficiente de regresión m , basados en análisis de rango, son los propuestos por Mood y Brown (1950), que aplican la estimación de ambos parámetros m y b simultáneamente usando la mediana estadística determinada mediante pruebas de ensayo y error. Adichie (1967) propone un método más restrictivo bajo el supuesto de que el conjunto de puntos, sobre el que se pretende ajustar la línea recta, sigue una función de distribución continua y simétrica con una función de densidad integrable, continua y cuadrada.

Theil (1950) propone un estimador muy simple para calcular m usando también la mediana estadística, mientras que Dytham (2011) y Sen (1968) estudian un estimador simple y robusto para el cálculo de m basado en el coeficiente de correlación por rangos del tau de Kendall (1955), donde un simple análisis no paramétrico puede usarse en lugar de una regresión normal.

El tau de Kendall es una medida estadística utilizada para medir la asociación entre dos medidas cuantitativas. Se trata de una medida de correlación por rangos que tiene en cuenta la similitud del ordenamiento de los datos cuando se clasifican por cada una de sus cantidades (Kendall, 1938; Kruskal, 1958). Dado un conjunto de puntos $(x_1, y_1)(x_2, y_2) \dots (x_n, y_n)$, cada par de puntos (x_i, y_i) y (x_j, y_j) se dice que son concordantes si $x_i > x_j$ y $y_i > y_j$, o si $x_i < x_j$ y $y_i < y_j$; discordantes si $x_i > x_j$ y $y_i < y_j$, o si $x_i < x_j$ y $y_i > y_j$; finalmente el par no es concordante ni discordante si $x_i = x_j$ o $y_i = y_j$. Con ello, el coeficiente tau de Kendall se define como sigue:

$$\tau = \frac{N_{pc} - N_{pd}}{\frac{n}{2}(n-1)} \quad (3.27)$$

siendo N_{pc} el número de pares concordantes, N_{pd} el número de pares discordantes y n el número total de pares. Dado que el denominador es el número total de pares, el coeficiente debe estar en el rango $-1 < \tau < 1$, siendo 1 cuando la concordancia entre las dos clasificaciones es perfecta (las dos clasificaciones son iguales), -1 cuando la

discordancia es máxima (una clasificación es la opuesta de la otra) y finalmente, en caso de que las variables x e y sean independientes, el coeficiente tiende a cero.

En base a lo anterior, se selecciona la mediana del conjunto de pendientes como el estimador para m y b , donde una pendiente simple se obtiene entre todos los posibles pares de píxeles i y j con coordenadas en la imagen (x_i, y_i) y (x_j, y_j) respectivamente. Finalmente, la pendiente seleccionada como la mejor estimación de m es la mediana de las pendientes tal y como se verá en el capítulo cinco sección 5.4.3.

Dado que la mayoría de píxeles que forman parte del conjunto de estudio pertenecen a la línea de cultivo, las pendientes en su mayoría pertenecen a pares de píxeles que pertenecen a dicha línea. Por este motivo, el estimador Theil-Sen es capaz de obtener la recta que mejor se aproxima a la línea de cultivo, sin que influyan los píxeles que puedan situarse en los límites más exteriores o que pertenezcan a otra línea de cultivo o a un rodal de malas hierbas. La figura 3.15 ilustra de forma esquemática esta situación.

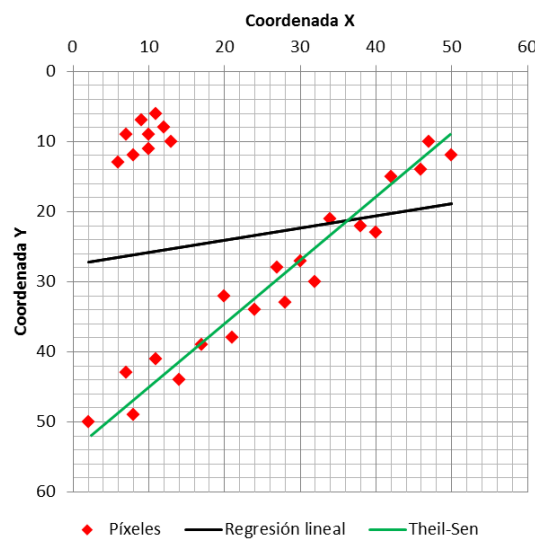


Figura 3.15 Comparativa de una recta obtenida por regresión lineal simple (línea negra) y otra recta obtenida utilizando el estimador Theil-Sen (línea verde).

Tal y como puede apreciarse en la figura 3.15, mientras la regresión lineal simple sí se ve influenciada por el conjunto de píxeles aislados, el estimador Theil-Sen es capaz de descartar dichos valores obteniendo una recta que se ajusta mejor a la línea de cultivo, lo que muestra la robustez que presenta este estimador a la hora de aproximar la recta que mejor atraviesa la nube de puntos con mayor densidad.

4

Segmentación de imágenes agrícolas en campos de maíz

4.1 Introducción

En AP, la identificación de plantas verdes es un proceso clave, ya que la mayoría de las aplicaciones requieren de esta tarea para diferentes propósitos. Es por ello que el éxito de los tratamientos específicos en AP, ya sean químicos, mecánicos o térmicos, dependen de la correcta identificación de las plantas verdes. En esta identificación juegan un papel importante varios factores entre los que destaca la intensidad luminosa, particularmente la procedente del sol como fuente de iluminación natural. Su estudio se aborda en la sección 4.2. En la sección 4.3 se estudia una técnica de segmentación y umbralización que constituye un paso crucial dentro del proceso de segmentación. Por último, un método de aprendizaje, basado en lo que se conoce como Máquinas de Vectores Soporte, se incluye en la sección 4.4.

4.2 Identificación y tratamiento de imágenes con alta intensidad luminosa

Los sistemas de visión por imágenes utilizados en tareas relacionadas con la agricultura y en particular en AP deben trabajar en condiciones meteorológicas adversas y en muchos casos variables en cortos espacios de tiempo. Uno de los problemas a los que se debe enfrentar es la iluminación, la cual afecta considerablemente a la calidad de la imagen, ya que una iluminación excesiva provoca una saturación física en el sensor

CCD haciendo que la imagen pierda calidad y en consecuencia que fallen los métodos basados en índices de vegetación para la extracción de los verdes. Esto se debe a que la mayoría de los píxeles pertenecientes al suelo o bien no son detectados, o bien se identifican como plantas verdes, dando lugar a falsos positivos. Cuando esto ocurre el sistema de visión está trabajando de forma inapropiada y necesita un nuevo ajuste del iris integrado en el sistema óptico o un ajuste en el tiempo de exposición si la cámara lo permite a través del correspondiente interfaz. Esto hace que sea necesario determinar si la calidad de la imagen es suficiente o no, con el fin de corregir posibles anomalías durante la adquisición de la misma.

A continuación se detallan los pasos propuestos para el tratamiento de este problema en la presente investigación, que suponen parte de la aportación principal de este trabajo de tesis y se recogen en Guerrero y col. (2012*b*).

4.2.1 Identificación de imágenes con alta intensidad luminosa

El objetivo final de la estrategia propuesta es determinar si una imagen dada es válida para la identificación de plantas verdes (cultivo y malas hierbas) o es inapropiada debido a sus altos niveles de intensidad producidos por exceso de iluminación. La estrategia propuesta se enmarca dentro de las técnicas de aprendizaje adaptadas a tal propósito, donde el objetivo consiste en tomar una decisión sobre la calidad de una imagen basándose en los niveles de intensidad de ésta. Si la calidad se considera suficiente, se puede aplicar un proceso de detección del verde para la identificación de plantas. Por el contrario, si la calidad de la imagen se considera insuficiente, la imagen deberá ser descartada o se deberá emitir una alarma indicando que el proceso de adquisición no está funcionando correctamente, siempre y cuando no sea posible realizar una corrección de los factores que provocan el fenómeno de iluminación no deseado. Algunas formas de corrección son la disminución del tiempo de exposición de la cámara mediante software o bien el cierre o apertura de forma automática del iris.

La estrategia que se propone para determinar la calidad de la imagen, en función de los niveles de intensidad, consta de una etapa de aprendizaje y otra de decisión. En la etapa de aprendizaje el objetivo es obtener el valor de umbralización que establece si la imagen contiene un exceso de intensidad o no. Con este propósito, se considera la iluminación como un factor decisivo y se aplica una transformación de la imagen de

datos original en el espacio de color RGB al espacio de color HSI, ambos descritos en el capítulo tres sección 3.2.1. Ésta es una práctica común para conocer, evaluar y corregir algunos efectos derivados de la iluminación (Cheng y col., 2001; Burks y col., 2000) y que normalmente proporciona resultados apropiados de segmentación. Por ejemplo en Tang y col. (2000) se utiliza un método para la segmentación de las plantas y del suelo utilizando un algoritmo genético basado en el modelo HSI.

La conversión del espacio de color RGB a HSI se realiza normalizando los valores RGB de la siguiente forma (Pajares y Cruz, 2007; Gonzalez y Woods, 2007):

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B} \quad (4.1)$$

Una vez normalizados, cada canal resultante en el modelo HSI se obtiene a través de las ecuaciones (4.2), (4.3) y (4.4).

$$H = \begin{cases} \theta & \text{si } B \leq G \quad H \in [0, \pi] \\ 2\pi - \theta & \text{si } B > G \quad H \in [\pi, 2\pi] \end{cases} \quad (4.2)$$

$$\theta = \cos^{-1} \left\{ \frac{1/2[(r-g) + (r-b)]}{\left[(r-g)^2 + (r-b)(g-b) \right]^{1/2}} \right\}$$

$$S = 1 - 3 \cdot [\min(r, g, b)] \quad S \in [0, 1] \quad (4.3)$$

$$I = \frac{1}{3}(r + g + b) \quad I \in [0, 1] \quad (4.4)$$

Por convenio, los valores H, S e I se convierten a valores dentro de los intervalos [0, 360], [0, 100] y [0, 255] respectivamente según las ecuaciones (4.5).

$$\begin{aligned} H &= H \times (180/\pi) \\ S &= S \times 100 \\ I &= I \times 255 \end{aligned} \quad (4.5)$$

En los experimentos que se han llevado a cabo durante la investigación presentada en esta memoria, parte de la conversión se ha omitido con el fin de disminuir el coste computacional de la transformación ya que el sistema de visión ha de trabajar en

tiempo real. Así, dado que para la estrategia propuesta no es necesario trabajar con los canales H y S, únicamente se realiza el cálculo para obtener la matriz de intensidad I, ecuación (4.4), y utilizar un valor de umbral en el intervalo $[0, 1]$ para identificar imágenes con altos valores de intensidad. Además, para disminuir el coste computacional, no se realiza la conversión de dicho canal a valores en el intervalo $[0, 255]$ que es útil a la hora de la representación de la imagen.

4.2.1.1 Fase de aprendizaje: ajustes del umbral

En un primer paso, se dispone de un conjunto de 250 imágenes, que se divide en dos subconjuntos, C1 y C2, figura 4.1. El procedimiento para asignar una imagen dada a cada conjunto se realiza como sigue:

- a) Se calcula el nivel de verde para cada imagen basándose en el índice de vegetación ExG y se aplica la umbralización utilizando el método de Otsu, ambos métodos explicados en el capítulo tres sección 3.3 y 3.4.2 respectivamente, tras lo cual se obtiene una imagen binaria.
- b) Un experto compara la imagen binaria con la imagen original teniendo en cuenta la situación real y el estado de las plantas en el cultivo. De este modo, si la imagen no se aproxima a lo que el experto considera una imagen válida, desde el punto de vista de un criterio agrícola, ésta se considera como inválida para un tratamiento específico por zonas.

De las 250 imágenes, 143 fueron clasificadas como válidas y por lo tanto poseen una calidad aceptable para la identificación de las plantas (cultivo y malas hierbas), y 107 fueron clasificadas como inválidas debido a los valores altos de intensidad que presentan. Las imágenes inválidas producidas por una iluminación excesiva, se clasificaron como pertenecientes al conjunto C1, determinándose que el porcentaje de verde (PG) de esas imágenes se sitúa siempre por debajo del 12% cuando en realidad dicho valor es claramente muy superior. Por el contrario, las imágenes válidas se asignaron al conjunto C2 y el porcentaje de verde que contienen es superior al 12%, por lo tanto, tras los diversos experimentos de prueba y error llevados a cabo, se determina que dicho porcentaje resulta ser un valor crítico en la distinción de ambos tipos de imágenes.

En el segundo paso, se transforma cada imagen del espacio de color RGB original al HSI y se obtiene el valor promedio de intensidad I del modelo HSI. En este paso se calcula el valor medio de intensidad para todas las imágenes pertenecientes al conjunto C1, cuyo valor resultante es 0.8769; del mismo modo se calcula el valor medio de intensidad de las imágenes del conjunto C2, obteniendo un valor de 0.3779. Como se quiere encontrar un umbral óptimo para la clasificación de una nueva imagen dada a cada conjunto, y considerando que las imágenes de intensidad pertenecientes a C1 tienen un mayor valor promedio que las de C2, se obtiene el valor mínimo de C1 y el valor máximo de C2 que se corresponden con los valores límite de cada conjunto. Estos valores son 0.6067 y 0.5881 respectivamente. Ambos valores se sitúan próximos a 0.6, por ello se selecciona este valor como el umbral para la toma de decisiones a la hora de determinar si una nueva imagen se considera o no con la calidad suficiente desde el punto de vista de la iluminación, es decir, si se trata de una imagen válida o inválida.

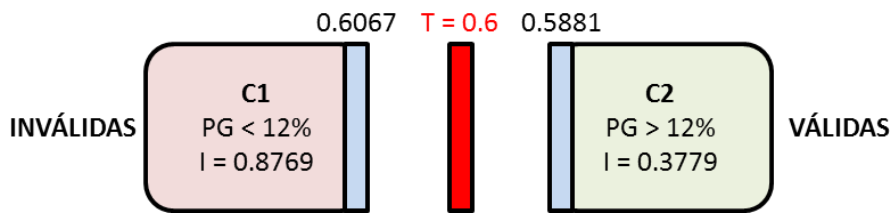


Figura 4.1 Esquema del proceso de aprendizaje. Conjuntos C1 y C2 con los datos característicos de cada uno, porcentaje de verde (PG), valor medio de intensidad (I) y umbral seleccionado (T).

4.2.1.2 Fase de decisión y procesamiento

Una vez se ha establecido el umbral, se aplica la siguiente estrategia para determinar si una nueva imagen de entrada se considera inválida por los valores altos de intensidad, o válida para la identificación de plantas verdes:

- Se transforma la imagen del espacio de color RGB a HSI.
- Se calcula el valor medio (m_I) de la intensidad, canal I en el modelo HSI.
- Si m_I es menor que un umbral T_I (fijado a 0.6 según el apartado anterior), la imagen se considera que tiene una calidad suficiente y es válida. En este caso la aplicación del índice de vegetación ExG (Woebbecke y col., 1995; Ribeiro y col., 2005), seguido de la estrategia de umbralización por el método de Otsu (1979) garantiza una extracción correcta de la parte verde.

En otro caso, se descarta la imagen y se genera una señal de alarma, al considerarse que la imagen fue adquirida de forma incorrecta y que por tanto es inválida para la extracción del verde.

La figura 4.2 muestra en detalle el proceso completo de decisión, desde la identificación de una imagen adquirida con una iluminación excesiva hasta cómo realizar la binarización para la identificación de las partes verdes en el caso de que los niveles de intensidad sean idóneos. Como se observa en dicho diagrama, existen dos partes claramente diferenciadas. La primera (parte izquierda), se encarga de identificar si una imagen es considerada válida para poder aplicar el proceso de binarización. Para ello, tras tomar la imagen se realiza la transformación del modelo RGB al HSI (punto a). Después se calcula el valor medio de intensidad para dicha imagen (punto b) y se compara con el umbral T_1 (punto c), fijado al valor 0.6 obtenido en la fase de entrenamiento según el procedimiento previamente descrito. Si la imagen no es válida no se puede procesar directamente, habría que descartarla para evitar el consumo de tiempo computacional o bien generar una alarma para indicar que el sistema no está funcionando correctamente e intentar corregir la captura de las sucesivas imágenes mediante el control del tiempo de exposición. Si esta acción de control no fuera posible, habría que abortar la misión para realizar un ajuste manual del iris del sistema óptico antes de proceder a mejorar la calidad de imagen aplicando técnicas tales como la corrección del *vignetting* y el balance de blancos, que se explicarán en la sección 4.2.2.1. Por el contrario, la segunda parte (derecha) del esquema de la figura 4.2 se realiza sólo cuando la imagen es apta para el proceso de binarización. Una vez se ha validado la imagen, se le aplica el índice vegetativo ExG, tras lo cual se obtiene la correspondiente imagen en escala de grises. Acto seguido se aplica el método de Otsu descrito en el capítulo tres sección 3.4.2, obteniendo una imagen binaria donde las partes blancas corresponden a las plantas y las negras al resto (tierra, piedras u otros elementos del suelo).

En este caso se ha decidido utilizar el índice de extracción de verde ExG debido a que en las imágenes utilizadas, la componente verde en las plantas domina sobre las otras dos componentes (roja y azul). Como además es superior a la componente verde que posee el suelo, se obtiene una imagen en escala de grises donde las plantas se manifiestan con una mayor relevancia en la escala de grises resultante. Por otro lado, se

ha aplicado el método de Otsu como umbralización porque es capaz de obtener un umbral dinámico válido para cualquier situación de luminosidad adecuada y además también se ha utilizado con buenos resultados en otras aplicaciones que requerían la umbralización de imágenes agrícolas (Meyer y Camargo-Neto, 2008; Guerrero y col., 2012a, b, 2013; Montalvo y col., 2012a, b, 2013a; Romeo y col., 2013a; Burgos-Artizzu y col., 2011; Guijarro y col., 2011, 2013).

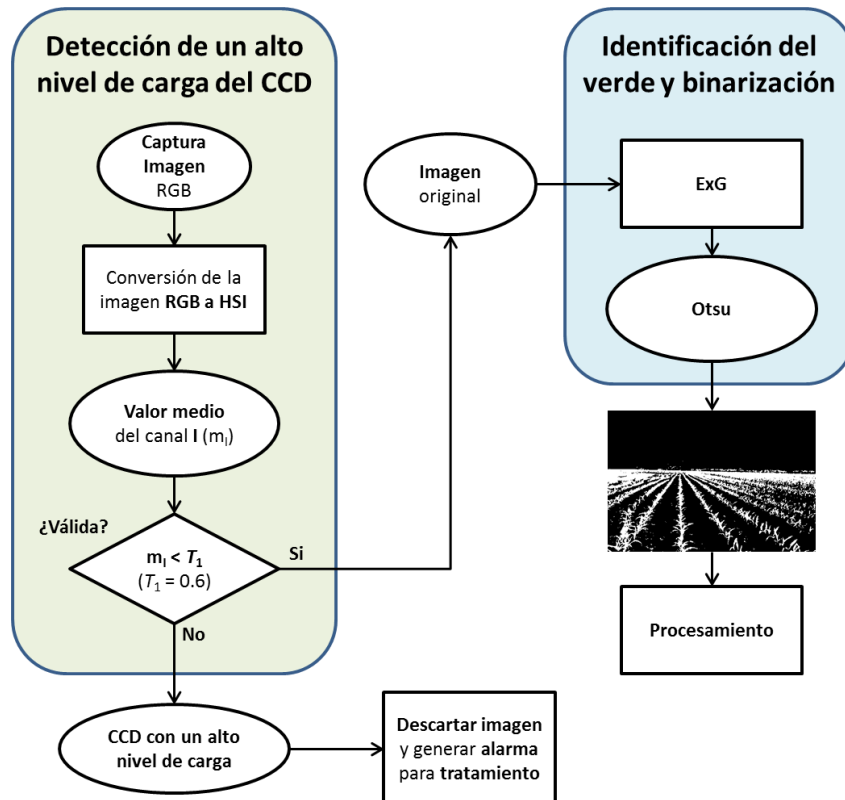


Figura 4.2 Esquema del proceso de decisión para determinar la validez de la imagen capturada y pasos en el proceso de binarización.

En la figura 4.4 se muestran los resultados obtenidos tras aplicar el proceso mostrado en el esquema de la figura 4.2 a partir de dos imágenes, una con niveles altos de intensidad, figura 4.3(a), y otra con niveles adecuados de intensidad, figura 4.3(b), ambas adquiridas con la cámara BASLER scA1400-17fc con una lente PENTAX C30811KP, cuyas especificaciones técnicas junto con las del sistema óptico asociado se describen en el capítulo seis sección 6.2.

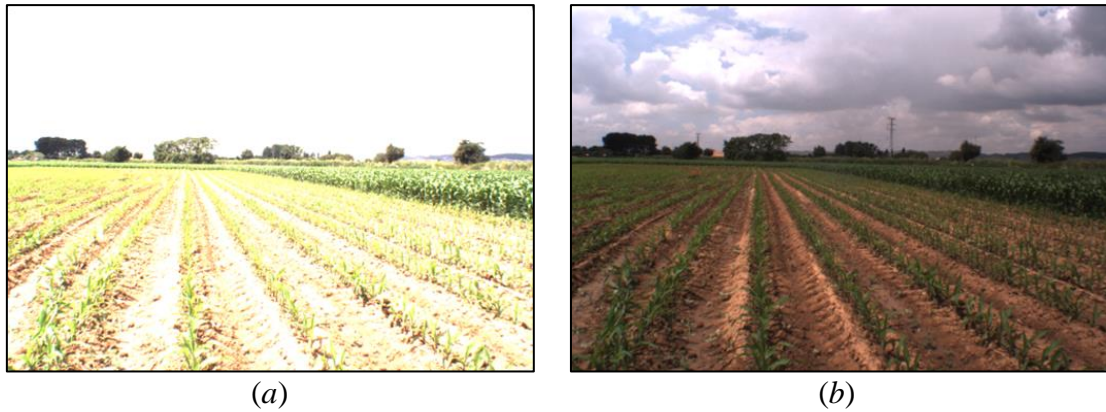


Figura 4.3 (a) Imagen original con un nivel alto de iluminación. (b) Imagen original con un nivel adecuado de iluminación.

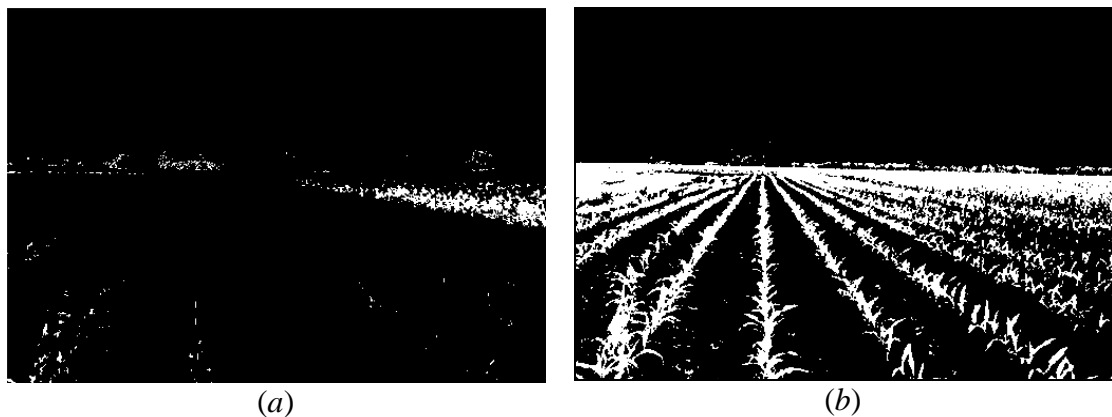


Figura 4.4 Imágenes obtenidas después de aplicar ExG y el método de Otsu. (a) A partir de la figura 4.3(a) con un alto nivel de intensidad, con valor medio de intensidad 0.876. (b) A partir de la figura 4.3(b) con un nivel adecuado de intensidad, con valor medio de intensidad 0.4688.

En el capítulo seis, sección 6.3.1 se analizan los resultados obtenidos con el fin de comprobar por un lado el porcentaje de éxito de la estrategia propuesta respecto de la obtención de los verdes y por otro el tiempo de proceso para su integración en sistemas que tienen que funcionar en tiempo real.

4.2.2 Tratamiento de imágenes

Una vez se determina que una imagen posee valores de intensidad excesivos, debido a un malfuncionamiento del sistema de visión, existen varias opciones antes de abortar la misión. En este punto se han diseñado dos métodos, uno basado en el análisis de blancos y otro en el análisis de la intensidad lumínica de la imagen. En ambos casos el fin es controlar automáticamente el tiempo de exposición del sensor CCD a la recepción de la reflectividad y la iluminación procedentes de la escena, permitiendo que el sistema se auto-ajuste sin necesidad de interrumpir el procesamiento de las imágenes.

También se ha diseñado una estrategia para corregir las anomalías de color existentes en una imagen si no fuera posible realizar un ajuste del tiempo de exposición, bien porque la cámara no lo permite y el ajuste del iris implica que las imágenes tomadas estén excesivamente oscuras, o bien porque el CCD está protegido mediante algún filtro que altera los colores de la imagen obtenida debido a cortes específicos en determinadas longitudes de onda. Concretamente, la estrategia contempla dos acciones, corrección de *vignetting* y balance de blancos

En ocasiones se puede realizar un ajuste automático de la apertura del iris siempre y cuando la cámara posea un sistema óptico que lo permita. Como se ha visto en el capítulo dos sección 2.3.2, existen sistemas ópticos cuya apertura del iris es ajustable automáticamente mediante un motor que responde a los cambios en los niveles de luz, aunque en situaciones con excesiva iluminación, la difracción y el desenfoque cuando la apertura del iris es demasiado pequeña puede ser problemática. Por este motivo existen lentes conocidas como de tipo P, que limitan el cierre del iris para evitar el desenfoque (difracción) que se produce cuando la apertura del iris es demasiado pequeña.

Pero no siempre se dispone de un sistema óptico con iris ajustable automáticamente. La situación más común es tener una óptica con iris de ajuste manual, que, por tanto, no podría ajustarse si el sistema está integrado en un robot autónomo. Por ello, se han diseñado varias alternativas a tener en cuenta dependiendo de si la cámara permite o no un ajuste en el tiempo de exposición (velocidad de obturación).

4.2.2.1 Cámaras con iris manual y que no permiten un ajuste del tiempo de exposición

Para este tipo de cámaras es necesario efectuar un ajuste sobre la propia imagen realizando transformaciones sobre el contenido en sus valores de intensidad. Es importante que el iris en este tipo de cámaras no esté excesivamente abierto, de este modo se puede obtener una imagen algo más oscura que con una apertura mayor pero que permite el procesamiento de la imagen de forma apropiada. Si el iris está excesivamente abierto, el sensor CCD recibe altas dosis de iluminación provocando imágenes con altos valores de intensidad, y en este caso habría que abortar la misión para realizar un ajuste del diafragma del iris de forma manual, ya que en tal caso se

produce una importante pérdida de información haciendo que este tipo de imágenes no sean de utilidad. Otro de los problemas que aparecen es el conocido como efecto *vignetting* en sus dos modalidades, mecánica y óptica, ambas descritas en el capítulo tres, sección 3.2.2. Para compensar y corregir los efectos adversos generados tanto por una iluminación deficiente como por el efecto *vignetting*, seguidamente se describen dos procedimientos de corrección.

a) Corrección del vignetting: para paliar este efecto negativo es necesario aplicar una corrección de intensidad en cada píxel sobre los tres canales espectrales de la imagen. La estrategia de corrección se basa en el efecto físico que genera el problema tal y como se explicó previamente, concretamente en el hecho de que los valores de intensidad de los píxeles, cuanto más próximos en distancia estén respecto al centro de la imagen menos será la variación de su valor por el efecto *vignetting*, y viceversa, siendo en las esquinas de la imagen donde se producirán las mayores variaciones. En la figura 4.5 se muestra una imagen que representa el patrón de variación y se observa cómo los valores situados en el centro de la imagen tienen un valor de 0 y los situados en las esquinas de 1, el resto de píxeles presentan valores en el rango $[0, 1]$ con una distribución en función de la distancia del punto al centro de la imagen. De esta forma se genera un filtro de corrección degradado capaz de minimizar los efectos del *vignetting*.

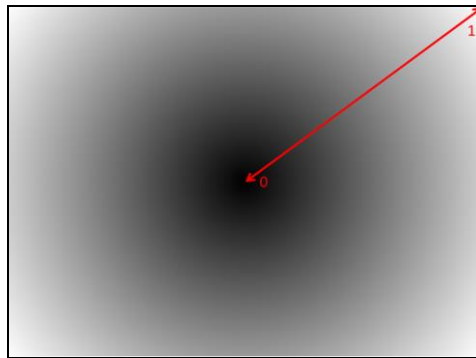


Figura 4.5 Filtro utilizado en la corrección del *vignetting*.

Para la obtención de dicho filtro se calculan las distancias relativas de cada píxel con respecto al centro de la imagen normalizadas en el intervalo $[0, 1]$ tal como se recoge en las ecuaciones (4.6) y (4.7).

$$dist(x, y) = \sqrt{(x - Centro_x)^2 + (y - Centro_y)^2} / distMax \quad (4.6)$$

$$distMax = \sqrt{(Centro_x)^2 + (Centro_y)^2} \quad (4.7)$$

siendo x e y las coordenadas del píxel a corregir, y $Centro_x$ y $Centro_y$ las coordenadas x e y del centro de la imagen. Una vez calculadas las distancias, cada canal espectral R , G y B se modifica independientemente según las expresiones (4.8).

$$\begin{aligned} R(x, y) &= 1 + FR \cdot dist(x, y) \\ G(x, y) &= 1 + FG \cdot dist(x, y) \\ B(x, y) &= 1 + FB \cdot dist(x, y) \end{aligned} \quad (4.8)$$

donde FR , FG y FB son factores de corrección constantes para cada canal.

De este modo, dependiendo de la causa que haya originado el efecto, pueden asignarse valores diferentes a cada componente con el fin de ajustar al máximo la corrección. En las pruebas realizadas con la cámara SVS4050CFLGEA incorporando el filtro Schneider UV/IR 486 (Schneider Kreuznach, 2014b), cuyas especificaciones técnicas junto con las del sistema óptico se describen en el capítulo seis sección 6.2, tras diversas pruebas de ensayo y error, se llegó a la conclusión de que los mejores resultados se obtienen con los siguientes factores de corrección para cada canal:

$$\begin{aligned} FR &= 0.3 \\ FG &= 0.0 \\ FB &= 0.0 \end{aligned} \quad (4.9)$$

Estos valores se justifican en base al hecho de que tal y como se ha mencionado en el capítulo tres sección 3.2.2, este filtro actúa directamente sobre las longitudes de onda del espectro en el infrarrojo, y dado que el rango de trabajo del filtro utilizado (Schneider UV/IR 486) tiene una pequeña tolerancia, ésta repercute también directamente sobre el canal rojo haciendo que los valores disminuyan con respecto a sus valores nominales, lo que ocasiona un predominio del verde sobre el rojo. Este filtro también realiza el corte de longitudes de onda en el ultravioleta, esto supone por efecto de su tolerancia en el otro extremo, que el canal azul se ve igualmente afectado como le ocurre al rojo. Ahora bien, debido a que nuestro problema se centra en la identificación de plantas verdes sobre suelo desnudo (compuesto por arena, arcilla y limo) con dominio del rojo, el problema se circunscribe a las componentes roja y verde, de forma que el efecto sobre la componente azul puede despreciarse. El efecto *vignetting*, como

se ha mencionado previamente, hace que predomine fuertemente el verde sobre el rojo, siendo tanto más acusado cuanto más lejos nos situemos del centro de la imagen, tal y como se muestra en la imagen de la figura 4.6(a). Para corregir esta alteración basta con aumentar los valores del canal rojo, sin necesidad de modificar el resto, lo que justifica los valores que aparecen en la expresión (4.9). La figura 4.6(b) muestra la imagen corregida mediante el procedimiento descrito y con los valores establecidos en (4.9). Como se puede apreciar en la figura 4.6(b), las esquinas y laterales de la imagen aparecen mucho más limpios, y el verde que antes se identificaba erróneamente como planta ha desaparecido en su mayoría dando lugar a una mejor identificación de la capa vegetal.

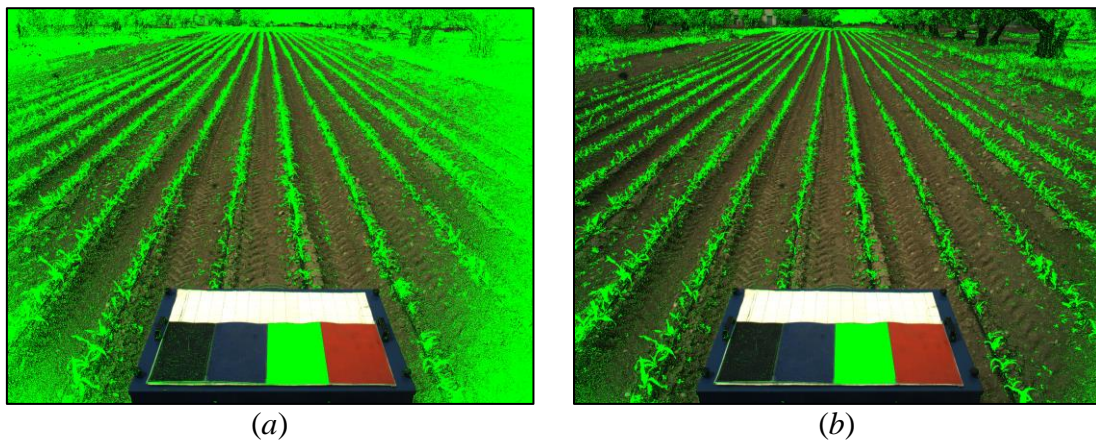


Figura 4.6 (a) Imagen donde se aprecia el *vignetting* en los laterales de la imagen y en las esquinas; (b) Imagen donde se aprecia la corrección del efecto *vignetting* al aplicar la corrección sobre el canal rojo.

b) Corrección de color mediante balance de blancos: una vez se ha corregido el *vignetting*, puede darse el caso de que los colores de la imagen sigan estando alterados, bien por el uso del algún filtro, como el ya mencionado UV/IR, o bien porque el diafragma del iris se encuentre excesivamente cerrado y la imagen no haya sido captada con valores de intensidad suficiente. Como se ha comentado anteriormente, una apertura excesiva del iris no sólo provoca *vignetting*, sino que el sensor de la cámara recibe demasiada luz y por lo tanto las imágenes capturadas aparecen con altos valores de intensidad en los tres canales espectrales, lo que impide la correcta identificación de las plantas.

Para corregir este tipo de anomalías en las imágenes se puede utilizar la técnica de balance de blancos, mediante la cual se escala toda la luminancia relativa de la imagen para que los objetos aparezcan con sus colores naturales a la vez que mantienen un cierto equilibrio entre ellos. Para ello, se considera una zona de referencia en la

imagen, la cual se sabe a ciencia cierta que corresponde a una superficie blanca en la escena. Se considera blanco el que aparece como tal para el ojo humano, sin entrar en detalles de matiz del color. La idea para la corrección es sencilla y se ilustra bien con el siguiente ejemplo: si la superficie blanca tiene un valor $G = 230$ y el valor de blanco se corresponde en la representación con el valor 255 para cada uno de los canales R , G y B , entonces los valores de G han de multiplicarse por el coeficiente $255/230$ para conseguir un blanco puro en la zona de referencia. De forma análoga se realiza el escalado para los otros dos canales, R y B , dando como resultado la siguiente matriz diagonal:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 255 / R_w & 0 & 0 \\ 0 & 255 / G_w & 0 \\ 0 & 0 & 255 / B_w \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (4.10)$$

donde R , G y B son los valores de las componentes roja, verde y azul ajustadas para un píxel; R' , G' y B' son las componentes roja, verde y azul del píxel en la misma posición espacial en la imagen original; y R_w , G_w y B_w son los valores de las componentes roja, verde y azul de un píxel o valor promedio sobre la zona blanca de referencia de cada componente de un conjunto de píxeles.

Para poder realizar el balance de blancos se requiere situar la región de referencia en alguna posición conocida en el tractor. De esta forma se conoce la parte de la imagen donde se proyectará tal región para obtener los valores R_w , G_w y B_w . En nuestros experimentos el panel de referencia está ubicado sobre la plataforma que alberga el sistema de energía por pilas de hidrógeno del tractor en la parte delantera del mismo. Las figuras 4.7 (a) y (b) muestran la mencionada plataforma.

Dado que el tractor está en continuo movimiento durante las labores agrícolas y además sometido a las vibraciones del motor, fijarse en un único punto blanco puede resultar un problema, ya no sólo a la hora de situarlo, sino que al estar trabajando en entornos de exterior puede darse el caso de que la luz incida de forma diferente sobre ese píxel originando que el resultado no sea el esperado. Por ello se toma como referencia una zona más amplia, 50x50 píxeles en el caso de la figura 4.7(a), en lugar de en un único píxel. De este modo se garantiza que intervengan varios valores de blanco en el cálculo de los valores promediados R_w , G_w y B_w , minimizando así los problemas derivados de la utilización de un único píxel.

En la figura 4.7(a) se muestra una imagen original sin balance de blancos, mientras que en la figura 4.7(b) aparece la misma imagen tras la aplicación del proceso de balance de blancos.

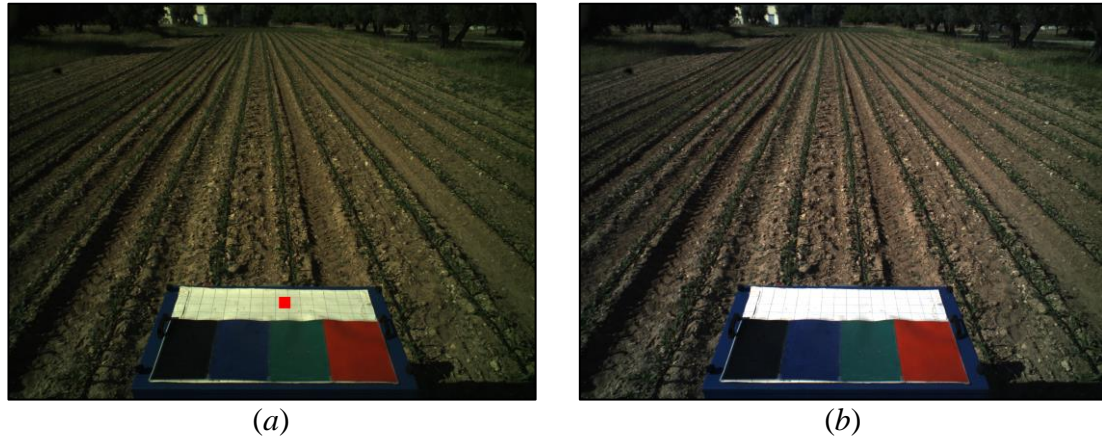


Figura 4.7 (a) Imagen original; (b) Imagen modificada con el balance de blancos.

En la imagen original el color natural está alterado en parte por el filtro Schneider UV/IR 486 instalado en el sistema óptico acoplado a la cámara SVS4050CFLGEA. Se puede apreciar cómo la imagen corregida mediante balance de blancos posee unos colores más equilibrados y menos verdosos. El cuadrado rojo que aparece en la imagen original, figura 4.7(a), corresponde a la zona que se toma como referencia para realizar el balance de blancos, que como puede apreciarse se sitúa sobre una zona más amplia que se sabe que es blanca. El ancho del rectángulo rojo se corresponde con las dimensiones indicadas de 50x50 píxeles de la imagen.

4.2.2.2 Cámaras que permiten un ajuste del tiempo de exposición

Para este tipo de cámaras existen diferentes métodos para ajustar el tiempo de exposición, bien basándose en el estudio de los tres canales de una imagen RGB (Montalvo y col. 2014, 2013b) o bien utilizando el canal Intensidad resultante de transformar la imagen del modelo original RGB al HSI (Guerrero y col., 2014). En este caso se han desarrollado dos procedimientos distintos, ambos con el fin de controlar el tiempo que el sensor CCD permanece expuesto a la entrada de radiación e iluminación externa. Uno basado en la técnica del balance de blancos descrita en el apartado anterior, y el otro basado en el estudio de una imagen con alta intensidad luminosa, y por lo tanto con exceso de iluminación, tal y como se ha visto en la sección 4.2.1. En la figura 4.8 se muestra un esquema del proceso general para el ajuste del tiempo de exposición.

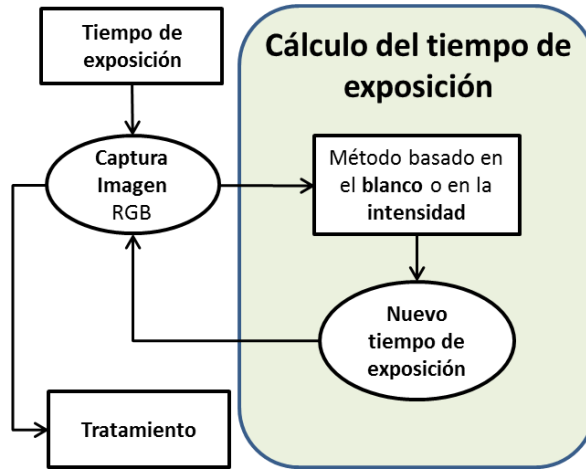


Figura 4.8 Esquema del proceso de ajuste del tiempo de exposición.

a) Ajuste del tiempo de exposición basado en el blanco: este método utiliza una zona de referencia blanca de características similares a la utilizada en el caso del balance de blancos. Esta zona se utiliza para identificar cuándo es necesario aumentar o disminuir el tiempo de exposición, es decir, hacer que el sensor CCD esté más o menos expuesto a la entrada de radiación e iluminación. Para ello se necesita determinar un valor de umbral superior para cada canal espectral R , G , B de la imagen, así se identifica cuándo el valor de blanco está próximo a la saturación con el fin de evitar que la imagen llegue a tener valores de intensidad elevados en cada uno de los canales. Dado que el nivel máximo que puede alcanzar un píxel blanco puro se corresponde con los máximos valores en los canales $(R = 255, G = 255, B = 255)$, se establecen tres valores de umbral U_R, U_G y U_B máximos para cada canal con el fin de identificar si el valor del tiempo de exposición ha de modificarse o no. El objetivo es obtener la diferencia de cada canal con respecto al umbral para evaluar si es necesario ajustar el tiempo de exposición o no, por ello se obtienen las diferencias para cada canal sobre la zona de referencia tal como se recoge en la expresión (4.11):

$$\begin{bmatrix} R_D \\ G_D \\ B_D \end{bmatrix} = \begin{bmatrix} U_R - R_w \\ U_G - G_w \\ U_B - B_w \end{bmatrix} \quad (4.11)$$

donde R_D , G_D y B_D son las diferencias de las componentes roja, verde y azul con respecto a los umbrales correspondientes para dicho canal (U_R, U_G, U_B) ; y R_w , G_w y B_w son los valores de la componente roja, verde y azul de un píxel, o valor promedio de

cada componente de un conjunto de píxeles, que pertenecen a una superficie blanca de la imagen original. En la práctica, dichos umbrales (U_R, U_G, U_B) se han fijado a 200 con el fin de mantener un tiempo de exposición aceptable de forma que el blanco se sitúe cerca de la saturación sin llegar a ella. Una vez se han obtenido las diferencias, cada canal se evalúa de la siguiente manera:

$$\begin{aligned} \text{si } C_D > 0 &\rightarrow C_V = 1 \\ \text{si } C_D < 0 &\rightarrow C_V = -1 \\ \text{en otro caso } &C_V = 0 \end{aligned} \quad (4.12)$$

donde C_D es la diferencia calculada anteriormente para cada uno de los canales (R_D, G_D, B_D) y C_V la variación que sufre cada canal (R_V, G_V, B_V) . Una vez obtenidas las variaciones de los tres canales, éstas se suman para comprobar si es necesario aumentar el tiempo de exposición o disminuirlo según las reglas (4.13).

$$\begin{aligned} \text{si } [R_V + G_V + B_V] > 1 &\rightarrow t(i) = t(i-1) + (t(i-1) / 2) \\ \text{si } [R_V + G_V + B_V] < -1 &\rightarrow t(i) = t(i-1) - (t(i-1) / 2) \end{aligned} \quad (4.13)$$

donde $t(i)$ es el tiempo de exposición actual calculado en función del valor del tiempo utilizado para la captura de la imagen anterior $t(i-1)$, siendo i el instante actual e $i-1$ el instante previo.

El procedimiento expresado en las ecuaciones (4.11), (4.12) y (4.13) se describe más detalladamente a continuación. Una vez obtenidas las diferencias de cada canal con respecto a sus umbrales, según la ecuación (4.11), a cada componente espectral R , G y B se le asigna un valor de variación, ecuación (4.12), +1 para una diferencia positiva que indica un aumento del tiempo de exposición, -1 para una diferencia negativa que indica un decremento del tiempo de exposición, o bien cero en cualquier otro caso para indicar que dicha componente no influye en la decisión. Posteriormente se suman las tres variaciones con el fin de tomar una decisión, ecuación (4.13), de esta forma siempre se tiene en cuenta el peso de las tres componentes, siendo el resultado aquel cuyo valor viene influenciado por al menos dos de las tres componentes en el mismo sentido, es decir, mínimo dos componentes indican que el tiempo de exposición debe aumentar, o bien indican que ha de disminuir. Lo que se pretende es mantener un tiempo de exposición con cierta estabilidad, para ello se necesita que al menos dos de las tres

componentes estén, o bien por encima o bien por debajo del umbral para llevar a cabo la corrección, ya que cuantos menos canales se tengan en cuenta, mayores correcciones se realizarán, pudiendo ocurrir que de una imagen a otra el tiempo de exposición varíe excesivamente, lo que de ninguna manera es deseable. Por ello, si el resultado de la suma de las tres variaciones de las componentes según la ecuación (4.13) resulta ser un valor positivo mayor que 1, significa que el nivel de blanco está por debajo del umbral fijado para los tres canales y por tanto es necesario incrementar el tiempo de exposición. Si por el contrario, el valor es negativo y menor que -1, el tiempo de exposición resulta excesivo lo que requiere una disminución del mismo. El intervalo del tiempo de exposición que se aumenta o disminuye viene dado por el tiempo de exposición anterior, siendo el valor medio de dicho tiempo el que habría que sumar o restar en función de los datos obtenidos. Con esto se consigue que el ajuste del tiempo de exposición sea proporcional al valor actual y que el cambio se realice de forma rápida y efectiva.

Uno de los problemas que plantea este método, es que cualquier mínimo cambio en la zona de referencia del blanco interfiere en el ajuste. Por ejemplo, si la zona se ensucia o una sombra se proyecta directamente sobre el blanco, el método llevará a realizar un ajuste sin tener en cuenta si la zona de referencia se sitúa o no en niveles aceptables para su procesamiento.

La figura 4.9 muestra dos imágenes consecutivas tomadas durante un giro del tractor. Como puede apreciarse, figura 4.9(a), cuando una sombra, provocada en este caso por la cabina del tractor donde está montada la cámara, invade la zona de referencia (cuadrado rojo), el método detecta una disminución en los niveles de blanco y por consiguiente aumenta el tiempo de exposición obteniendo acto seguido la imagen mostrada en la figura 4.9(b). Esto se debe a que los valores obtenidos en la ecuación (4.11) eran todos positivos. El resultado produce una sobreexposición del sensor generando la imagen mostrada, cuya calidad es a todas luces muy deficiente, particularmente en la parte del campo identificada como zona de interés donde se han de distinguir los verdes.

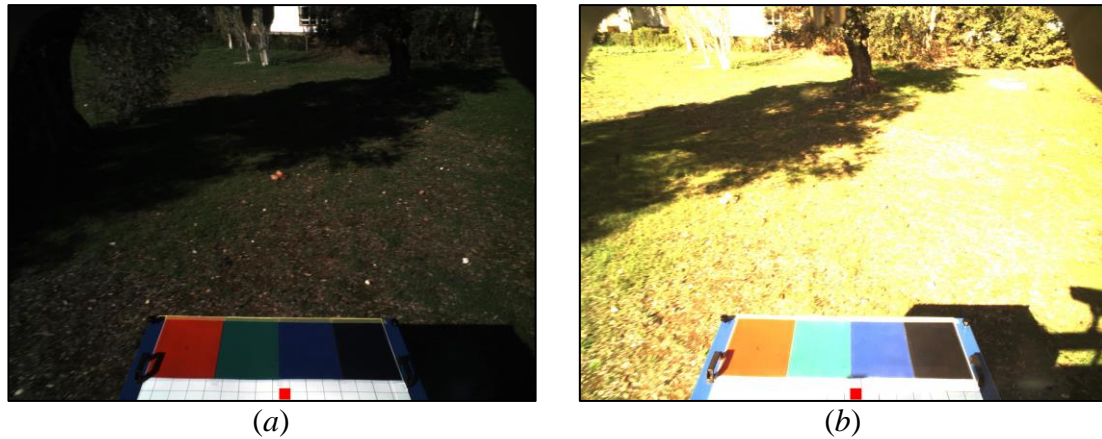


Figura 4.9 Imágenes consecutivas durante el movimiento del tractor autónomo. (a) Imagen con un tiempo de exposición bien ajustado. (b) Imagen con un tiempo de exposición excesivo.

La mencionada zona de interés (ROI, *Region Of Interest* en inglés) en nuestro caso es aquella zona de la imagen que comprende la parte del campo real donde se realizará la identificación de las líneas de cultivo, así como de las malas hierbas. Tal y como se verá con más detalle en el capítulo seis, en los experimentos llevados a cabo durante la investigación relacionada con el proyecto RHEA (2014), esta zona, situada sobre el suelo, se extiende a lo largo de 2 m desde una distancia de 3.3 m de la base del plano vertical que incluye el soporte donde va montada la cámara, y hasta una distancia de 5.3 m desde la base del mismo plano, figura 4.10.

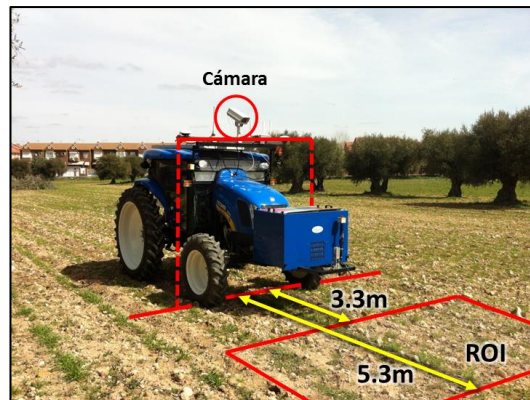


Figura 4.10 Imagen donde puede verse la zona de interés (ROI).

En la figura 4.11(a) se muestra la ROI desde otra perspectiva, se observa que esta zona abarca el ancho de cuatro líneas de cultivo, y dado que están espaciadas 0.75 m entre ellas, la anchura de la zona hace un total de 3 m. Trabajar en una zona de 3 m (4 líneas de cultivo) es un requerimiento impuesto por el implemento mecánico/térmico desarrollado en el proyecto RHEA y con el que se han realizado parte

de los experimentos descritos en la presente memoria. En cualquier caso, en lo que respecta a los métodos diseñados para el procesamiento de imágenes, tal y como se verá en el capítulo seis sección 6.5, la configuración del sistema de visión es modificable a través de un archivo de configuración, permitiendo a posteriori asumir cualquier cambio de equipamiento (cámara e implementos) o su modificación, aumentando o disminuyendo el número de líneas, su longitud y por consiguiente las dimensiones de la ROI. La ROI, vista desde la cámara instalada en el tractor, es la que se muestra en la figura 4.11(b), donde se puede apreciar que esta zona se sitúa siempre en la parte central de la imagen.

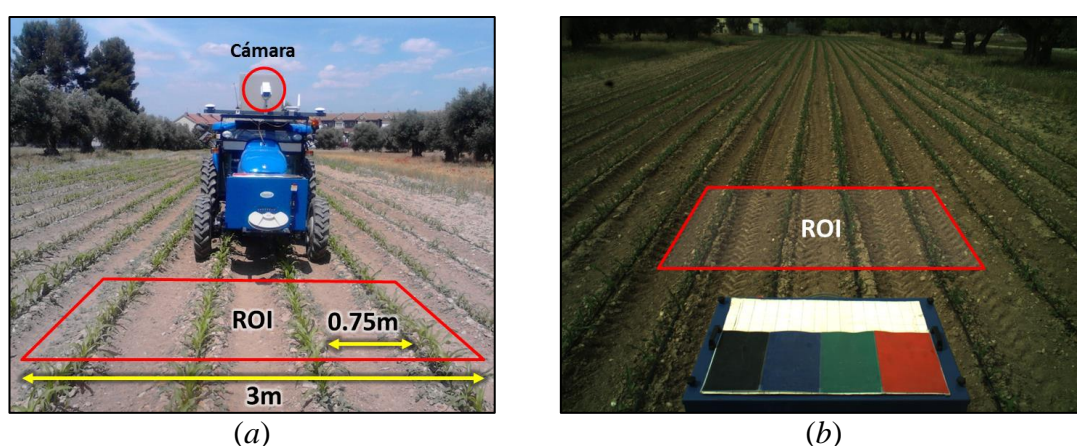


Figura 4.11 (a) Imagen donde puede verse la zona de interés (ROI) delante del tractor. (b) Imagen tomada de la misma escena a través de la cámara montada en el tractor.

El problema descrito en esta sección suele aparecer en entornos de exterior donde los dispositivos tienen un emplazamiento fijo y las condiciones de iluminación pueden afectar al punto que se toma como referencia para el ajuste. La estrategia propuesta resulta útil en situaciones en las que la iluminación de la zona de referencia no está afectada por elementos incontrolados (sombras del tractor) u oclusiones. Por ello, este procedimiento de ajuste no resulta apto para situaciones en las que la zona de referencia puede verse afectada por elementos ajenos al propio sistema.

Con el fin de resolver esta problemática, para la segunda implementación se ha utilizado el conocimiento adquirido en la fase experimental descrita en la sección 4.2.1.1 para la detección de imágenes con alta intensidad luminosa y se ha aplicado al procedimiento descrito en Guerrero y col. (2014).

b) Ajuste del tiempo de exposición basado en la intensidad lumínica: tal y como se propone en Guerrero y col. (2014), se estudia el comportamiento de los valores de intensidad del canal I obtenido al transformar la imagen del modelo original RGB al HSI. Esta transformación se aplica en la imagen sobre la ROI situada varios metros por delante del tractor y por tanto no afectada, en lo que a iluminación se refiere, por elementos propios del tractor, como puede ser la sombra de la cabina. En cualquier caso, incluso estando afectada por alguna adversidad externa, tal como sombras de la cabina del tractor o de árboles cercanos, al ser la zona donde se ha de realizar el tratamiento de la imagen, cualquier corrección sobre ella resulta siempre favorable. Una vez obtenido el valor medio de intensidad, si éste es inferior a un determinado umbral T_1 , a definir más adelante, se incrementa el valor del tiempo de exposición, y si es superior a un segundo umbral T_2 , también a definir posteriormente, se disminuye el tiempo de exposición. El procedimiento se sintetiza como sigue:

- a) Transformar la imagen del espacio de color RGB a HSI.
- b) Calcular el valor medio de la intensidad, m_I (canal I en el modelo HSI) de la zona de interés (ROI).
- c)
$$\begin{aligned} \text{si } m_I < T_1 &\rightarrow t(i) = t(i-1) + (t(i-1) / 2) \\ \text{si } m_I > T_2 &\rightarrow t(i) = t(i-1) - (t(i-1) / 2) \end{aligned}$$

siendo $t(i)$ el tiempo de exposición en el instante i calculado en función del tiempo de exposición en el instante anterior $t(i-1)$.

Los umbrales T_1 y T_2 se han fijado a 0.1 y 0.4 respectivamente. Estos valores se han seleccionado para evitar llegar al límite de saturación establecido a 0.6 en la sección 4.2.1.1 mediante el razonamiento allí expuesto. Como en el caso anterior, el valor del tiempo de exposición que se incrementa o disminuye viene dado por el tiempo de exposición anterior, sumando o restando el valor medio de dicho tiempo, que como antes, se calcula con el fin de que el ajuste del tiempo de exposición sea proporcional al valor actual con el que se están adquiriendo las imágenes, garantizando así que el ajuste se realizará de forma rápida y efectiva.

4.2.3 Proceso completo para el tratamiento de imágenes

A continuación se muestra el nuevo esquema del proceso propuesto, figura 4.12, tras añadir al proceso de identificación de imágenes con altos valores de intensidad, descrito en la figura 4.2, las nuevas funcionalidades y características. Como puede observarse, si la imagen no alcanza valores de intensidad altos (superiores a 0.6) se considera apta para el proceso de binarización tal y como ocurría anteriormente, pero ahora el valor T_1 pasa a valer 0.4 con el fin de no alcanzar el máximo nivel de saturación fijado previamente a 0.6 en la sección 4.2.1.1.

Una vez la imagen se considera válida al tener valores de intensidad razonables, se comprueba si la imagen tiene un valor medio de intensidad por debajo del cual se considera que no tiene la calidad suficiente, ya que puede darse el caso de que por una mala configuración del iris (diafragma excesivamente cerrado) o un fallo de la cámara, la imagen obtenida sea negra o prácticamente negra ($m_I \simeq 0$). Por ello, en el supuesto de que el valor medio de intensidad m_I sea inferior a un umbral T_3 fijado a 0.05, la imagen no se considerará con calidad suficiente para su procesamiento. Como se ha indicado previamente, esto supone un malfuncionamiento del sistema de visión que, en el caso de que la imagen sea demasiado oscura, supone que el sensor CCD ha sido impactado con un nivel bajo de intensidad y dependiendo del modelo de cámara o sistema óptico cabe la posibilidad de realizar todavía diferentes acciones con el fin de continuar con el procesamiento automático: a) si permite un ajuste automático del diafragma del iris no es necesario abortar la misión, bastará con descartar esa imagen y ajustar el iris para que las sucesivas imágenes se tomen correctamente; b) si la cámara permite un ajuste del tiempo de exposición, como en el caso anterior, no es necesario abortar la misión y bastará con descartar esa imagen y ajustar el tiempo para mejorar las sucesivas capturas. En el supuesto de que cualquiera de las acciones anteriores no sean factibles o la cámara haya dejado de funcionar, bien por falta de comunicación con el procesador, por sobrecalentamiento u otro problema, será necesario abortar la misión para realizar un ajuste manual del diafragma del iris o solucionar el malfuncionamiento de la cámara.

Si por el contrario, la imagen posee un valor medio de intensidad m_I superior al umbral T_3 , el subsiguiente proceso depende de si la cámara o el sistema óptico lleva asociado algún tipo de filtro de corte UV/IR que pueda provocar *vignetting*. En caso afirmativo, se realiza una corrección del *vignetting* seguido de un ajuste del balance de blancos con el fin de mantener el equilibrio entre los colores que se han visto modificados. Si por el contrario el sistema no posee filtro, pero el valor medio de intensidad m_I se sitúa por debajo del umbral T_2 fijado a 0.1 de acuerdo con los experimentos descritos la sección 4.2.2.2, también es necesario realizar un ajuste del balance de blancos con el fin de corregir los valores de intensidad en los tres canales espectrales para obtener una imagen de calidad. Es decir, un valor medio de intensidad m_I tal que $T_3(0.05) < m_I < T_2(0.1)$ implica que la imagen ha de ser corregida en el sentido indicado aplicando la técnica del balance de blancos. Por lo tanto, si la imagen final se ha modificado por alguno de estos motivos, será la modificada la que se utilizará para la binarización (extracción de verdes y umbralización), mientras que si no existe filtro y la imagen posee un valor medio de intensidad m_I tal que $T_2(0.1) < m_I < T_1(0.4)$ la imagen utilizada en el proceso de binarización será la original tal cual fue tomada con la cámara sin ningún tipo de ajuste ni corrección. En otro caso, si $m_I \geq T_1(0.4)$ o $m_I < T_3(0.05)$, será necesario realizar un ajuste sobre el tiempo de exposición debido a un exceso o a una falta de iluminación respectivamente.

Por otro lado, si la imagen posee valores de intensidad altos o bien una mala configuración o un fallo de la cámara genera una imagen altamente saturada con resultado de una imagen blanca o prácticamente blanca ($m_I \approx 1$) durante sucesivas capturas, al igual que ocurre en el caso contrario cuando la imagen es negra o prácticamente negra, será necesario abortar la captura de imágenes y por tanto la misión.

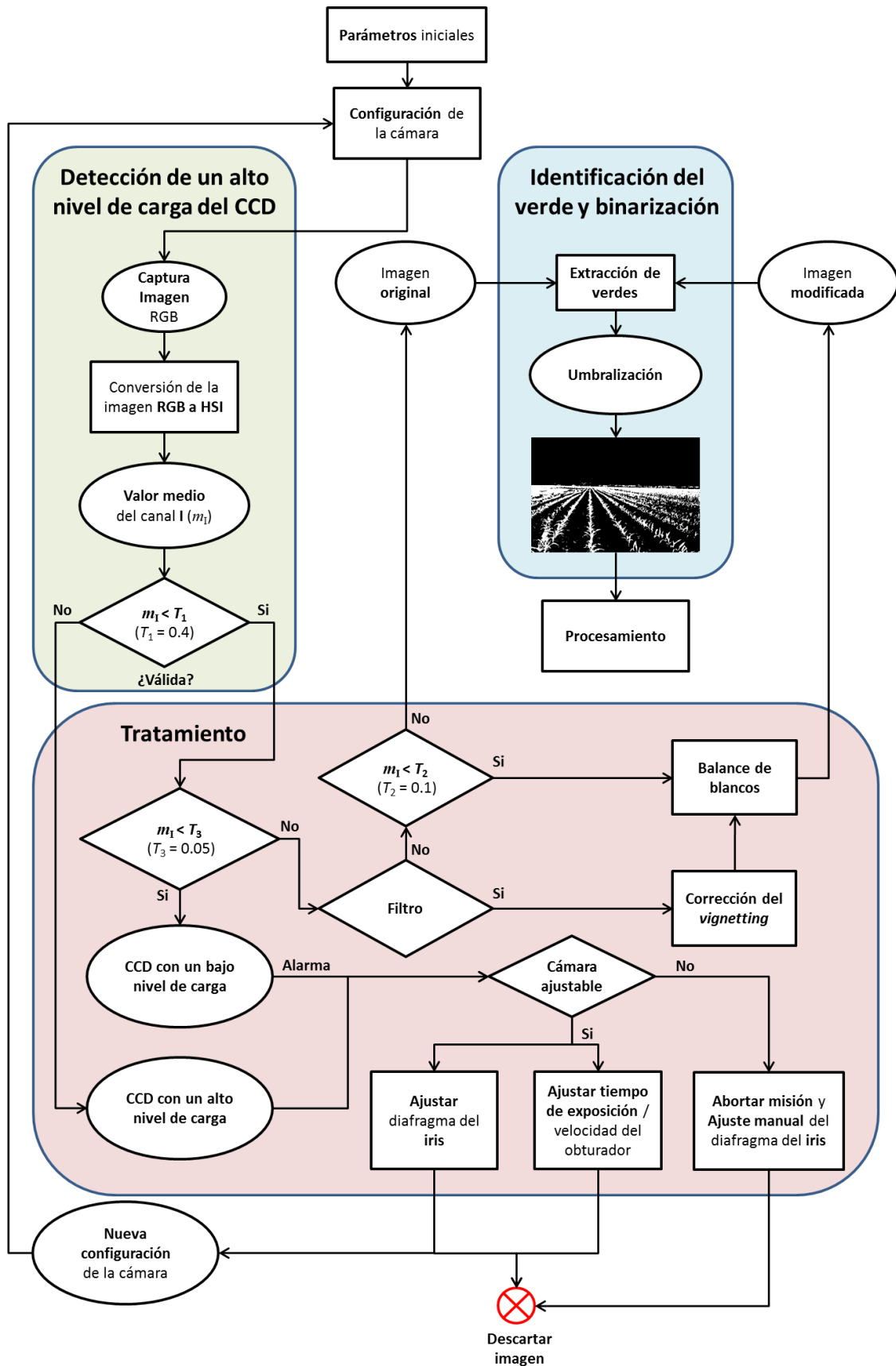


Figura 4.12 Esquema completo para la identificación de imágenes alta o insuficientemente iluminadas, su corrección, tratamiento y binarización para el posterior procesado.

El esquema que se propone se basa en un diseño de arquitectura abierta, de forma que en cualquier momento es posible reemplazar cualquiera de los módulos existentes por otro que realice funciones similares o de otra naturaleza. Esto permite incluir nuevas técnicas en el futuro, sin necesidad de variaciones sustanciales en el diseño, a medida que la investigación avance y de cara a la obtención de mejores resultados. Tal es el caso del módulo denominado “Identificación de verde y binarización”, que puede ser susceptible de cambios, tanto de cada una de sus partes (extracción de verdes, sección 3.3; umbralización, sección 3.4) como del módulo completo (secciones 4.3 y 4.4). Con el diseño del esquema propuesto y las técnicas desarrolladas se garantiza que la entrada a dicho módulo ha sido controlada y que la imagen de partida posee unos colores equilibrados, con unos niveles de intensidad válidos para un posterior procesamiento, tras la corrección de los eventuales problemas provocados por el uso de filtros o por una deficiente iluminación de la escena.

4.3 Identificación de plantas y binarización basada en $YD_B D_R$

Las técnicas descritas anteriormente, balance de blancos, corrección del *vignetting* y ajuste del tiempo de exposición, se utilizan como paso previo a la segmentación de la imagen. Sin ellas no se podrían obtener imágenes de partida con la calidad suficiente, y por consiguiente buenos resultados en la segmentación.

Con el fin de utilizar otros modelos de color que difieran del RGB a la hora de extraer los verdes con la subsiguiente binarización, se ha diseñado un nuevo método basado en $YD_B D_R$. Como se ha descrito en el capítulo tres sección 3.2.1.4, el modelo $YD_B D_R$ se basa en una componente de luminancia (Y) que representa el brillo general de la imagen, y dos componentes de cromaticidad (D_B y D_R) relacionadas directamente con el color, D_B con la componente azul y D_R con la componente roja. Tal y como se ha comentado anteriormente, el problema principal a resolver es la obtención de dos conjuntos de píxeles, uno asociado a los píxeles que representan las plantas verdes y otro con el resto de elementos que aparecen en el fondo, que en su mayoría tendrán una dominancia de la componente espectral verde y roja respectivamente. Es en este punto donde la cromaticidad D_R adquiere un papel relevante a la hora de identificar las partes verdes para discriminarlas del resto.

La conversión de una imagen del modelo RGB al $YD_B D_R$ (Shi y Sun, 2000) se realiza aplicando la matriz de transformación (4.14).

$$\begin{bmatrix} Y \\ D_B \\ D_R \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.450 & -0.883 & 1.333 \\ -1.333 & 1.116 & 0.217 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4.14)$$

donde R , G y B son las componentes espectrales roja, verde y azul para un determinado píxel en la imagen original representada en el modelo de color RGB e Y , D_B y D_R son los mencionados valores la luminancia y crominancias para dicho píxel en el modelo $YD_B D_R$. En las figuras 4.13 se muestran en escala de grises las tres componentes Y , D_B y D_R a partir de una imagen RGB.

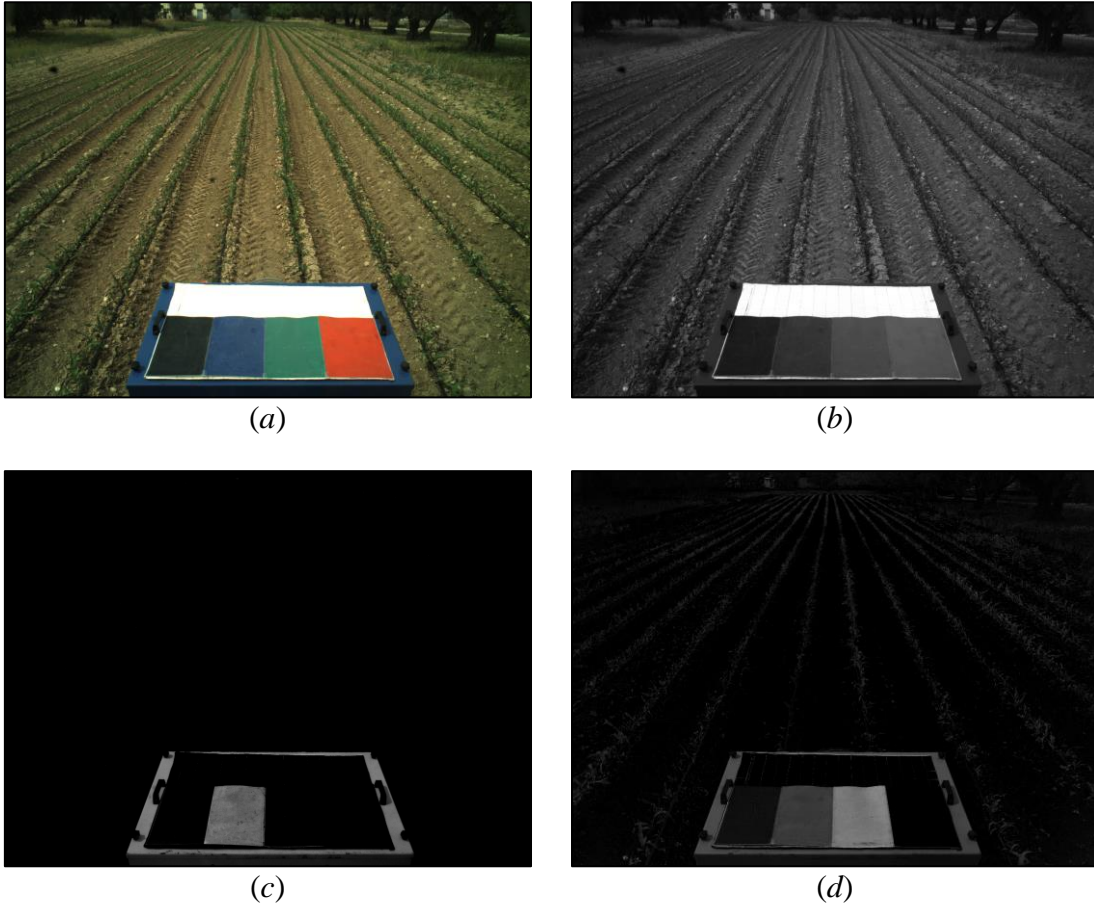


Figura 4.13 Descomposición de una imagen RGB en $YD_B D_R$ (a) Imagen original; (b) Imagen de la luminancia (Y); (c) Imagen de crominancia azul (D_B); (d) Imagen de crominancia roja (D_R).

Tomando como referencia el panel de colores mostrado en la parte inferior de la imagen original de la figura 4.13(a), se puede observar que la crominancia D_B , figura 4.13(c), está relacionada únicamente con la componente azul, mientras que la crominancia D_R está ligada a la componente roja y verde, figura 4.13(d), de forma que cuanto mayor es la componente verde de un píxel en la imagen original, con mayor intensidad aparece ese píxel en la imagen resultante transformada, y a su vez, cuanto mayor es la componente roja de un determinado píxel en la imagen original menor es el valor de ese píxel en la imagen resultante. Por ello, la crominancia D_R resulta apropiada para distinguir tonalidades verdosas, que son las que abundan en los píxeles de las plantas a identificar (cultivo y malas hierbas), a la vez que las tonalidades rojizas que son las que abundan en los píxeles del suelo. Por tanto, se puede concluir que la componente D_R está directamente relacionada con las tonalidades del suelo y de las plantas, mientras que la crominancia D_B lo está con el cielo.

Una vez identificada la crominancia D_R como la componente más apropiada, es necesario encontrar el umbral a partir del cual se puede realizar la distinción entre plantas y suelo. Para ello, se utiliza una base de imágenes compuesta por 4800 imágenes originales bajo diferentes estados de crecimiento del cultivo de maíz, distintos grados de iluminación y distintos niveles de infestación de malas hierbas. Dado que en la componente D_R aparece una clara influencia de la componente azul, como puede apreciarse en la figura 4.13(d), es importante descartarla en el cálculo del umbral. Para ello, a cada una de las imágenes se le resta la componente azul de la componente roja, obtenidas a partir de las respectivas crominancias siguiendo las expresiones (4.15) y (4.16).

$$\begin{aligned} D_B(x, y) &= B(x, y) - Y(x, y) \Rightarrow B(x, y) = D_B(x, y) + Y(x, y) \\ D_R(x, y) &= R(x, y) - Y(x, y) \Rightarrow R(x, y) = D_R(x, y) + Y(x, y) \end{aligned} \quad (4.15)$$

$$\begin{aligned} R(x, y) - B(x, y) &= (D_R(x, y) + Y(x, y)) - (D_B(x, y) + Y(x, y)) = \dots \\ \dots &= D_R(x, y) - D_B(x, y) = D'_R(x, y) \end{aligned} \quad (4.16)$$

siendo x e y las coordenadas del píxel. Con esto se obtiene una nueva imagen (D'_R) en escala de grises, que es la imagen de partida para realizar el proceso de binarización.

Con el fin de mejorar el tiempo de procesado, en lugar de aplicar el método de Otsu a cada imagen para la obtención de un umbral automático para la binarización, se lleva a cabo una fase de aprendizaje con un total de 2000 imágenes a las que se aplica el método de Otsu descrito en el capítulo tres sección 3.4.2 en la componente D_R , obteniendo como promedio un valor de umbral de 7 en el intervalo $[-340, 340]$, que se corresponde a un valor de 130 en el intervalo $[0, 255]$ de RGB para todo el conjunto de imágenes. Una vez obtenido dicho valor de umbral, el método de extracción de verdes y binarización basado en $YD_B D_R$ se resume como sigue:

- a) Obtener la componente D_R resultado de la transformación de la imagen desde el espacio de color RGB al espacio de color $YD_B D_R$.
- b) Si $D_R(x, y) > T (= 7) \Rightarrow D_R(x, y) \in Plantas$
en caso contrario $D_R(x, y) \in Suelo$

Es decir, todos los píxeles con valores de crominancia D_R que superen el umbral T fijado a 7, obtenido durante la fase de aprendizaje, se consideran pertenecientes a una planta, mientras que los que no superen dicho umbral se consideran como pertenecientes al suelo u otros elementos.

Como puede apreciarse en la figura 4.14, es importante destacar el hecho de que utilizando sólo la crominancia D_R , todo lo que sea identificado como planta puede también pertenecer al cielo u otras zonas azules, a partes oscuras tales como sombras, piedras, mangueras de riego, etc., o bien a partes claras como puede ser un brillo o una zona altamente iluminada, es decir, tan sólo se elimina lo que sea componente exclusivamente perteneciente al suelo.

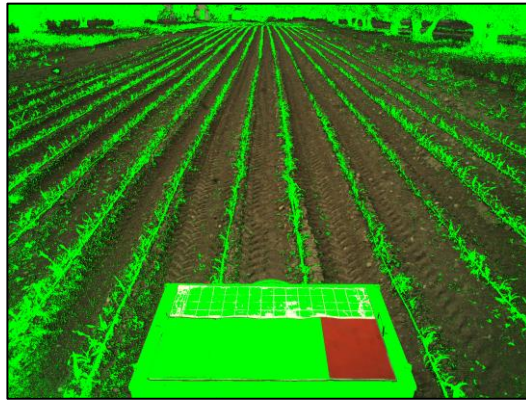


Figura 4.14 Imagen resultante al aplicar el nuevo método basado en $YD_B D_R$.

Por ello, sobre la imagen RGB es importante realizar un filtrado previo, donde aquellos píxeles cuya proximidad al negro o al blanco sea pequeña sean excluidos en el proceso de umbralización, siendo a la vez identificados directamente como otros elementos con el fin de que no sean considerados como plantas e induzcan a errores. Para su identificación, tras los diversos experimentos de ensayo y error en los que se seleccionaron distintas zonas con predominancia del negro y blanco, se llega a la conclusión de que aquellos píxeles cuya suma de las tres componentes espectrales RGB esté por debajo de 120 o por encima de 500, se excluyen del proceso.

También, de forma análoga habría que descartar aquellos píxeles que pertenecen al cielo o partes azules (como ocurre en la imagen reseñada con el morro del tractor). Para ello se debe realizar un filtrado similar a los anteriores sobre la imagen RGB, o bien como en este caso utilizar la cromaticidad D_B , ya que tal y como se ha puesto de manifiesto en la figura 4.13(c) identifica correctamente las zonas azules de la imagen.

La figura 4.15 muestra el resultado de aplicar el proceso de filtrado anterior. Ahora únicamente aparecen identificadas en su mayoría las plantas pertenecientes al maíz, pero hay que destacar el hecho de que, por ejemplo en la esquina inferior izquierda, algunas plantas no son identificadas correctamente debido a que esta zona de la imagen no ha recibido la luz suficiente como para superar el nivel inferior de umbral, con valor 120, establecido previamente.

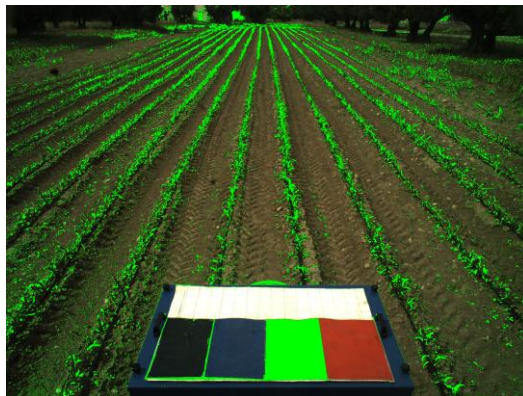


Figura 4.15 Imagen resultante al aplicar los filtros al nuevo método basado en $YD_B D_R$.

4.4 Segmentación e identificación de plantas enmascaradas y sin enmascarar mediante Máquinas de Vectores Soporte

Como norma general, para llevar a cabo la identificación del verde y la binarización de una imagen se elige un tratamiento basado en píxeles con la premisa de que las plantas pertenecientes al cultivo y a las malas hierbas tienen una componente espectral verde cuyos valores son superiores a los correspondientes valores de las respectivas componentes roja y azul. Como se ha indicado anteriormente, un método que se ha usado con buenos resultados en aplicaciones agrícolas, es la extracción del verde mediante ExG para transformar la imagen a escala de grises y la posterior umbralización de la imagen mediante el método de Otsu, que obtiene un umbral óptimo a partir del histograma. Ahora bien, esta combinación no siempre proporciona los mejores resultados, especialmente cuando se trabaja con imágenes donde el color verde de las plantas se ve alterado, ya sea porque el cultivo ha sido tratado con el fin de eliminar malas hierbas y las plantas se están secando, o porque parte del cultivo está enmascarado por los sedimentos tras una fuerte lluvia o como consecuencia de un riego reciente.

4.4.1 Análisis de índices vegetativos

La segmentación de la imagen se centra en la separación de plantas verdes (cultivo y mala hierba) del resto de elementos (suelo, piedras, rastrojo y otros). Tanto las plantas pertenecientes a las líneas de cultivo como las malas hierbas poseen unas componentes espectrales de color similares y durante el periodo de tratamiento el nivel de desarrollo es muy similar. Por esta razón, resulta relativamente complicado distinguirlas mediante técnicas basadas en sistemas de visión con discriminación, por ejemplo, por altura (Chen y col., 2013; Kise y Zhang, 2008; Kise y col., 2005; Rovira-Más y col., 2008). Recientemente se han realizado distintos experimentos en campos de maíz basados en sistemas LIDAR (*Light Detection And Ranging*) en Andújar y col. (2013) para distinguir entre cultivo, mala hierba y el suelo. Por otro lado, Andújar y col. (2011) han utilizado sensores opto-electrónicos basados en la emisión de luz propia, operando en las bandas del rojo e infrarrojo cercano, cuyo fundamento se centra en la discriminación entre la vegetación y el suelo debido a la distinta respuesta de las plantas frente al suelo, mientras que Zaman y col. (2011) han desarrollado un sistema basado en

micro-cámaras instaladas en la parte frontal del vehículo. En el caso de los sensores opto-electrónicos y de las micro-cámaras la idea consiste en localizar zonas de malas hierbas situadas entre las líneas, de forma que los sensores no alcanzan el cultivo. La discriminación se realiza por diferencias entre vegetación y suelo, por tanto se trata de dos sistemas con un diseño muy diferente al utilizado en esta investigación, que consta de una única cámara a bordo del tractor y cubriendo una zona más amplia que incluye plantas de cultivo, malas hierbas y suelo desnudo. Por las razones anteriores, en relación a nuestra investigación, la mejor opción para identificar malas hierbas y cultivo es la aplicación de índices de vegetación sobre las imágenes obtenidas con el sistema de visión instalado en el tractor autónomo.

Los índices de vegetación son métodos muy probados en aplicaciones agrícolas, como se ha visto en el capítulo tres apartado 3.3. Guijarro y col. (2011) proponen una combinación de índices de vegetación, que se elige y utiliza en este trabajo debido a su probado buen funcionamiento en campos de maíz. Es importante resaltar que los índices de vegetación en su mayoría identifican las plantas siempre y cuando no estén ocultas por los sedimentos y sean claramente diferenciables. Para tratar con esta situación que aparece con cierta frecuencia en un escenario real se propone el método que se describe a continuación.

4.4.2 Planteamiento del enfoque

Esta estrategia está diseñada para la identificación de plantas verdes (cultivo y mala hierba) en campos de maíz para tratamientos específicos por zonas, ya sean mediante productos químicos, control térmico o manipulación mecánica donde el objetivo final es la eliminación de malas hierbas. El maíz es un cultivo de regadío y sin protección frente a precipitaciones. Después de una lluvia fuerte o tras una sesión de regadío, el material del suelo se mezcla con el agua, de forma que el barro generado impregna la capa vegetal, principalmente aquellas partes situadas junto al suelo o muy próximas a él. En este caso, la componente espectral verde, dominante en las plantas de forma general, se ve enmascarada por la componente espectral roja que es la dominante en los materiales existentes en el suelo. La figura 4.16(a) muestra una imagen donde se puede apreciar este efecto en la zona central derecha cerca del borde inferior de la imagen, y también en el extremo final de las hojas de maíz que están orientadas hacia el suelo. La figura 4.16(b) muestra el mismo efecto en la parte central izquierda.

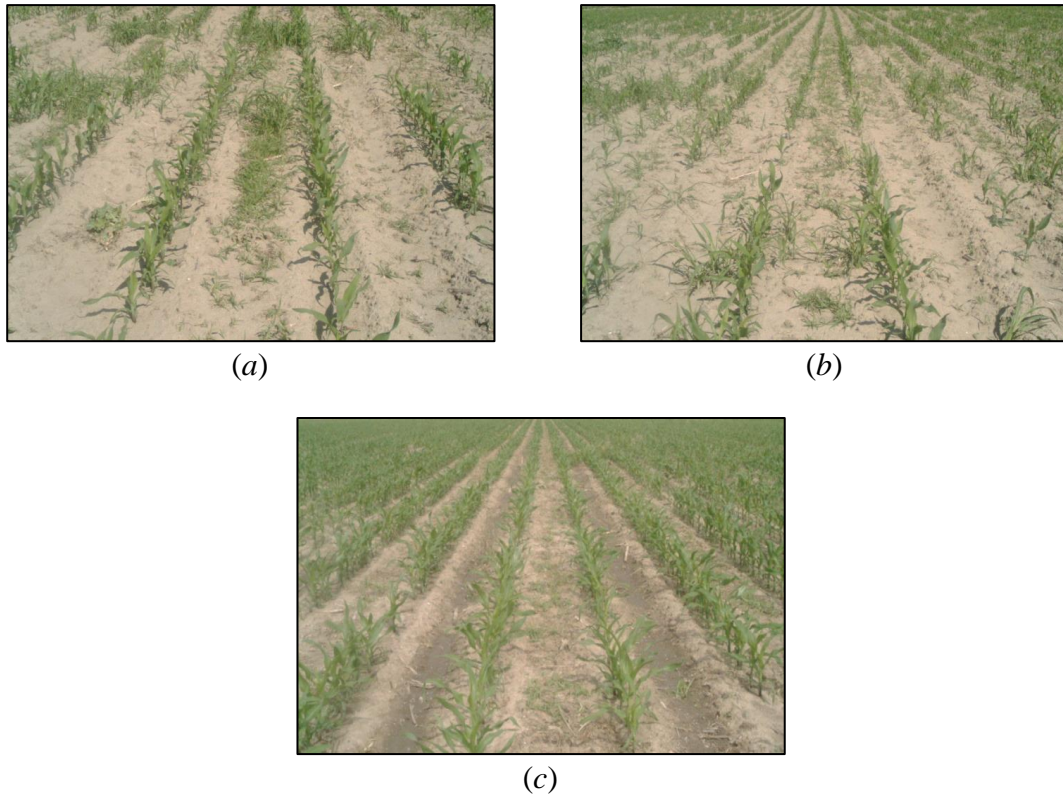


Figura 4.16 (a) Las malas hierbas de la zona central inferior y de la zona derecha aparecen como enmascaradas. (b) Las malas hierbas de la zona central aparecen enmascaradas. (c) El herbicida fue aplicado dos días antes (las malas hierbas en la zona central de las líneas de cultivo avanzan hacia una etapa de secado), además, el campo también recibió una lluvia directa.

Estas imágenes fueron adquiridas tras un riego natural producido por lluvia. En un caso como éste, los métodos de identificación del verde o capa vegetal, basados en el cálculo de índices vegetativos, fallan. De hecho, el suelo y las plantas enmascaradas por el barro se identifican como suelo cuando se aplica un índice para la extracción de verde. Esto puede dar lugar a un problema importante al identificar las densidades de malas hierbas de forma incorrecta.

Con el fin de solucionar el problema previamente descrito, se propone una estrategia (Guerrero y col., 2012a) basada en las Máquinas de Vectores Soporte (*Support Vector Machines*, SVM). Este método es lo suficientemente robusto y funciona correctamente bajo diferentes condiciones de iluminación provocadas por días nublados o soleados (Tian y Slaughter, 1998). Además ha demostrado, como se verá más adelante, ser válido para identificar plantas tanto con una componente espectral verde enmascarada como sin enmascarar, por lo que es un método adecuado para las evaluaciones que se realizan post-tratamiento del campo. De hecho, cuando las malas

hierbas han sido tratadas, comienza un proceso de secado a medida que van muriendo; esto significa que pierden parte de su color verde dominante por la desaparición paulatina de la función clorofílica, en cuyo caso la componente espectral dominante pasa a ser la roja. Mediante esta técnica también se puede evaluar un posible daño en el cultivo siempre que se produzca un comportamiento similar al descrito con respecto a las componentes verde y roja. La figura 4.16(c) muestra una serie de rodales de malas hierbas en la zona central de las líneas de cultivo que ya aparecen con cierto grado de deshidratación después de un tratamiento químico con herbicida aplicado dos días antes de la captura de la imagen. El método propuesto funciona adecuadamente ante ambas situaciones de pérdida de verdor por parte de las plantas. La eficacia del método diseñado permite verificar su viabilidad para tareas autónomas en agricultura basándose en el procesamiento de imágenes aún bajo las circunstancias adversas especificadas. Por simplicidad, de aquí en adelante, tanto las plantas enmascaradas como las que están en proceso de secarse se denominarán indistintamente plantas enmascaradas. Las SVM constan de una fase de aprendizaje seguida de otra de decisión.

4.4.3 Fase de aprendizaje basada en Máquinas de Vectores Soporte

La fase de aprendizaje se compone de varios procesos que se muestran en el esquema de la figura 4.17, y que se agrupan en tres tareas principales:

1. Identificación del verde seguido de una umbralización mediante el método de Otsu para obtener una imagen binaria, la cual se divide en dos clases.
2. Identificación de los vectores soporte (*support vectors*) asociados a cada clase, basado en SVM.
3. Cálculo de los valores medios de cada conjunto de vectores soporte, los cuales establecen los límites para la separación entre las clases con plantas enmascaradas y sin enmascarar.

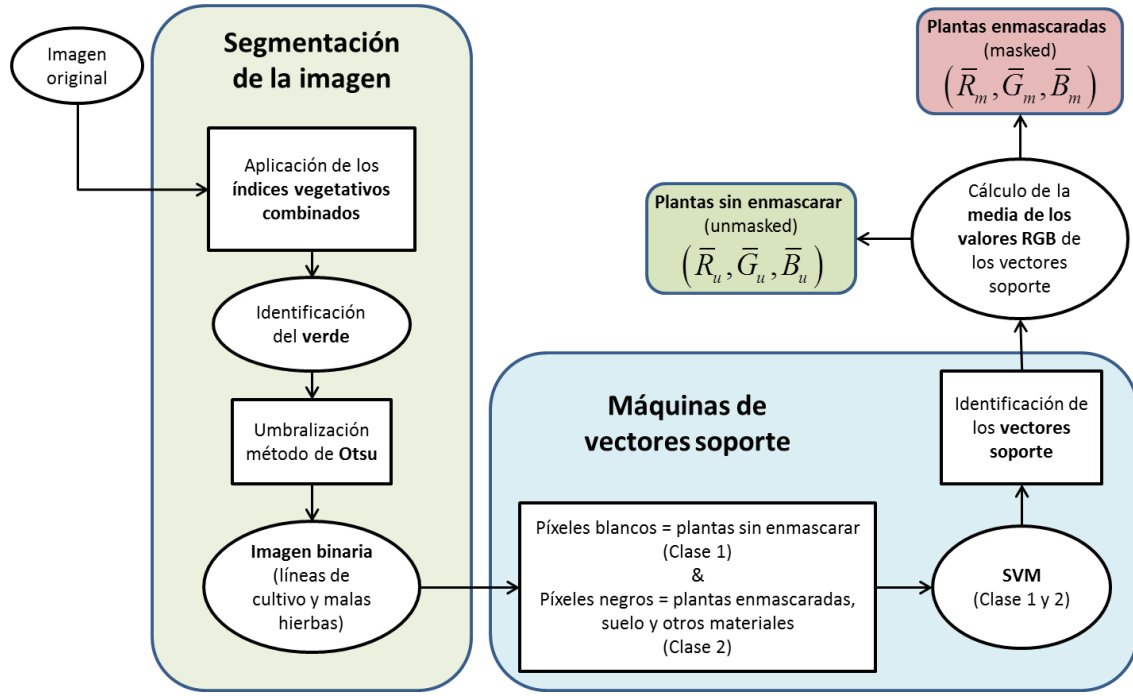


Figura 4.17 Esquema del proceso de aprendizaje.

a) Identificación del verde y binarización: dada una imagen original de entrada en el espacio de color RGB, y tomando como base la combinación de índices propuesta en Guijarro y col. (2011) se combinan los índices ExG, CIVE y VEG para obtener el valor COM tal y como se describe en el capítulo tres sección 3.3, que se simplifica siguiendo la expresión (4.17).

$$COM = w_{ExG}ExG + w_{CIVE}CIVE + w_{VEG}VEG \quad (4.17)$$

donde w_{ExG} , w_{CIVE} y w_{VEG} son los pesos para cada índice, representando su importancia relativa en la combinación. Guijarro y col. (2011) establece los valores para los tres pesos participantes en la combinación, pero en nuestro caso el término relacionado con el índice $ExGR$ ha sido excluido debido a que las sombras se identifican erróneamente. A tenor de la exclusión, los pesos para el resto de los índices se han recalculado de forma proporcional a los valores de partida, quedando finalmente de la siguiente forma: $w_{ExG} = 0.36$, $w_{CIVE} = 0.47$ y $w_{VEG} = 0.17$.

Los valores resultantes de la combinación COM se escalan linealmente dentro del intervalo $[0, 1]$ y posteriormente este resultado se umbraliza aplicando el método de Otsu, obteniendo una imagen binaria donde los píxeles blancos están asociados a plantas sin enmascarar. Por el contrario, los píxeles negros se identifican en la imagen

original como pertenecientes a plantas enmascaradas, suelo y otros materiales presentes en el campo.

b) *Identificación de los vectores soporte*: basándose en la imagen binaria, se crea una partición con dos clases. La Clase 1 (C_1) contiene aquellos píxeles que provienen de la imagen original y han sido marcados como píxeles blancos en la imagen binaria. Para cualquier píxel blanco situado en una posición espacial conocida, se obtiene el valor de sus tres componentes espectrales en la imagen original, es decir $x = (R, G, B) \in C_1$. De forma similar la Clase 2 (C_2) contiene aquellos píxeles que han sido marcados como negros en la imagen binaria. Por lo tanto, C_1 y C_2 pueden considerarse como dos subconjuntos que contienen patrones tridimensionales pertenecientes a dos clases diferentes. El número total de patrones es n , es decir, $n = \text{card}(C_1) + \text{card}(C_2)$. La salida del sistema SVM es $y \in \{+1, -1\}$, donde $+1$ y -1 están asociados a las clases C_1 y C_2 respectivamente. El conjunto de entrenamiento se construye como (x_k, y_k) , $k = 1, \dots, n$; donde cada x_k pertenece a cualquiera de las clases C_1 o C_2 e y_k denota la etiqueta a la que pertenece. El objetivo del proceso de entrenamiento basado en SVM es encontrar una función de decisión que separe los datos tridimensionales en ambas clases. Como se aprecia en la figura 4.18, los patrones de las diferentes muestras x_k , sobre la imagen tridimensional aparecen solapados.

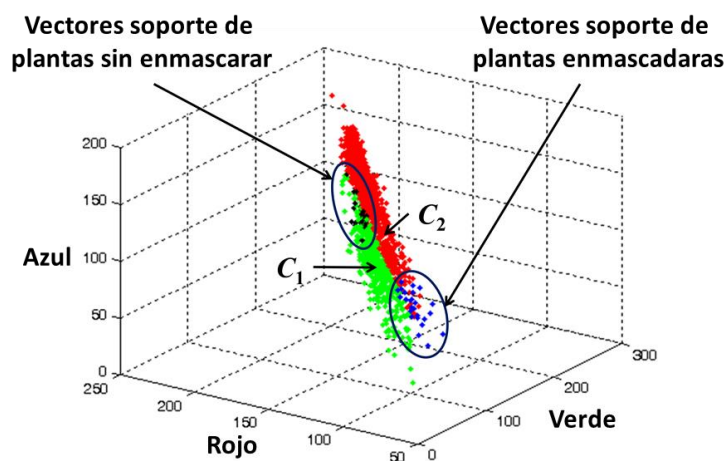


Figura 4.18 Distribución de patrones y vectores soporte en el espacio tridimensional RGB.

Transformando los vectores de entrada a un espacio multidimensional mediante funciones de transformación no lineales, el método SVM posee la capacidad de

encontrar la función de decisión incluso existiendo datos solapados (Cherkassky y Mulier, 1998; Vapnik, 2000). La forma general de la función de decisión SVM viene dada por la expresión (4.18).

$$f(x) = \sum_{k=1}^n \alpha_k y_k H(x_k, x) - b \quad (4.18)$$

donde x es un vector tridimensional genérico cuyas componentes son los tres valores espectrales R , G , B de la imagen original; b es un sesgo cuyo valor constante se establece más adelante. La ecuación (4.18) define una función $f(x)$ como una combinación lineal con el núcleo H , centrado en cada punto de datos. La elección del núcleo utilizado en los experimentos realizados se justifica y se describe en el capítulo seis sección 6.3.5. Los parámetros α_k con $k=1, \dots, n$ en (4.18) son la solución para el siguiente problema de optimización cuadrática:

$$Q(\alpha) = \alpha_k - \frac{1}{2} \sum_{k,h=1}^n \alpha_k \alpha_h y_k y_h H(x_k, x_h) \quad (4.19)$$

Bajo las siguientes restricciones,

$$\sum_{k=1}^n \alpha_k y_k = 0 \quad \text{y} \quad 0 \leq \alpha_k \leq \frac{C}{2}, \quad k=1, \dots, n \quad (4.20)$$

donde H es el producto escalar del núcleo mencionado anteriormente, y C es un parámetro de normalización. En Cherkassky y Mulier (1998) se menciona que no existe un buen desarrollo teórico sobre cómo seleccionar el mejor valor de C . En numerosas aplicaciones que involucran los SVMs se asigna un valor constante relativamente grande, como 2000, que es el que finalmente se ha utilizado en este trabajo. Los patrones x_k de datos, asociados con valores α_k distintos de cero se denominan *vectores soporte* (*support vectors*), éstos se renombran como s_k por comodidad y constituyen el fundamento de la investigación realizada para la separación de los dos tipos de plantas mencionados, es decir plantas verdes enmascaradas y sin enmascarar, en lugar de utilizar la función de decisión, que es lo habitual por norma general. Los vectores soporte son los puntos de datos con la mayor información disponible. La distancia mínima de los hiperplanos de separación entre las clases a los puntos de datos se conoce como margen τ (Cherkassky y Mulier, 1998). Un hiperplano de separación es óptimo

si el margen es máximo. La distancia entre un plano de separación y un patrón x_k dado, es $y_k |f(x_k)|/\|w\|$, donde w es un vector perpendicular al hiperplano definido por la ecuación (4.21).

$$w = \sum_{k=1}^n \alpha_k y_k x_k \quad (4.21)$$

El problema de encontrar el hiperplano óptimo es el de encontrar aquella w que maximiza τ . Existen infinitas soluciones que difieren sólo en la escala de w , de modo que para limitar las soluciones se fija la escala en el producto de τ y el vector normal de w , expresión (4.22).

$$\tau = \frac{1}{\|w\|} \quad (4.22)$$

donde $\|\cdot\|$ es una norma, siendo elegida la norma euclídea en los experimentos realizados en el presente trabajo de investigación. El parámetro de sesgo b mencionado en la ecuación (4.18) se calcula a partir de la expresión (4.23).

$$b = -\frac{1}{s} \sum_{k=1}^s (ws_k - y_k) \quad (4.23)$$

donde s_k son los s vectores soporte, es decir, los patrones de datos s asociados a α_k distintos de cero; y_k se inicializa con respecto al patrón que le corresponde en cada clase $(+1, -1)$.

c) Descripción de las clases: una vez se obtienen los vectores soporte, se dispone de dos conjuntos de vectores soporte, S_1 y S_2 , asociados a las clases C_1 y C_2 respectivamente, donde $S_1 \equiv \{s_k^1\}$ con $k = 1, \dots, n_1$ y $S_2 \equiv \{s_k^2\}$ con $k = 1, \dots, n_2$ y n_1 al igual que n_2 representan el número de vectores soportes pertenecientes a cada clase. Es bien conocido en el contexto de las SVM, que los vectores soporte representan los patrones más significativos de la clase a la que pertenecen. Estos vectores caen justo en el margen de separación de las clases, es decir, se sitúan cerca del hiperplano, por lo que los vectores soporte de S_1 están cerca de los patrones de la clase C_2 y viceversa. Por ello, los vectores soporte establecen los límites entre las clases. De este modo, los

vectores soporte de S_2 son exactamente los patrones más cercanos a la clase C_1 . Como la clase C_1 representa las plantas con un cierto grado de verde, los vectores soporte de S_2 son los patrones con el mayor grado de verde de la clase C_2 . Estos patrones a priori han sido asignados a una clase diferente a la que contiene las plantas verdes, pero ahora son identificados como patrones con cierto grado de verde, debido a su similitud con los patrones de la clase C_1 . Bajo esta suposición los vectores soporte de la clase C_2 se identifican como plantas enmascaradas o plantas afectadas por el tratamiento, que ya han comenzado el proceso de deshidratación. Se pueden clasificar todos los vectores soporte de S_2 a través del cálculo de dos medidas estadísticas, media $\bar{x}_m = (\bar{R}_m, \bar{G}_m, \bar{B}_m)$ y desviación estándar $\sigma_m = (\sigma_{mR}, \sigma_{mG}, \sigma_{mB})$. Del mismo modo se procede con los vectores soporte de C_1 , media $\bar{x}_u = (\bar{R}_u, \bar{G}_u, \bar{B}_u)$ y desviación estándar $\sigma_u = (\sigma_{uR}, \sigma_{uG}, \sigma_{uB})$ donde los subíndices m y u denotan plantas enmascaradas (*masked*) y sin enmascarar (*unmasked*) respectivamente. La media y la desviación estándar anteriores se obtienen en la fase de aprendizaje, figura 4.17.

4.4.4 Fase de decisión

Una vez se han identificado las dos clases mediante su media y desviación estándar, éstas se usan posteriormente en el proceso de toma de decisiones.

Dada una nueva imagen que contiene plantas enmascaradas y sin enmascarar, el objetivo es identificar los píxeles que pertenecen a cada tipo de planta. Como las plantas se identifican por su componente espectral verde, se calcula el porcentaje de verde de \bar{x}_m y \bar{x}_u según la expresión (4.24).

$$r_{mG} = \bar{G}_m / (\bar{R}_m + \bar{G}_m + \bar{B}_m) \text{ y } r_{uG} = \bar{G}_u / (\bar{R}_u + \bar{G}_u + \bar{B}_u) \quad (4.24)$$

También se calcula la desviación relativa con respecto a la componente espectral verde a partir de la expresión (4.25).

$$t_{mG} = \sigma_{mG} / \bar{G}_m \text{ y } t_{uG} = \sigma_{uG} / \bar{G}_u \quad (4.25)$$

Finalmente, dado un píxel de la nueva imagen con sus correspondientes componentes, $x = (R, G, B)$, éste se clasifica con arreglo a las reglas de decisión (4.26).

$$\text{Regla 1: } x \in C_2 \text{ si } r_{mG} - t_{mG} \leq G / (R + G + B) \leq r_{uG} - t_{uG} \quad (4.26)$$

$$\text{Regla 2: } x \in C_1 \text{ si } r_{uG} - t_{uG} < G / (R + G + B)$$

Estas reglas establecen los límites del verdor. La Regla 1 define una región de verdor entre los vectores soporte S_1 y S_2 la cual permite identificar los píxeles pertenecientes a plantas enmascaradas o afectadas por el tratamiento. La Regla 2 define una región con los mayores valores del índice de verdor, es decir, plantas sin enmascarar.

Las figuras 4.19 (a) y (b) muestran la imagen COM y el resultado después de aplicar Otsu, ambas obtenidas a partir de la imagen original de la figura 4.16(a), que es una imagen representativa del conjunto usado para el aprendizaje. Como se ha indicado previamente, la imagen corresponde al cultivo tras sufrir una lluvia directa donde el barro generado ha impregnado parte de las hojas de maíz.

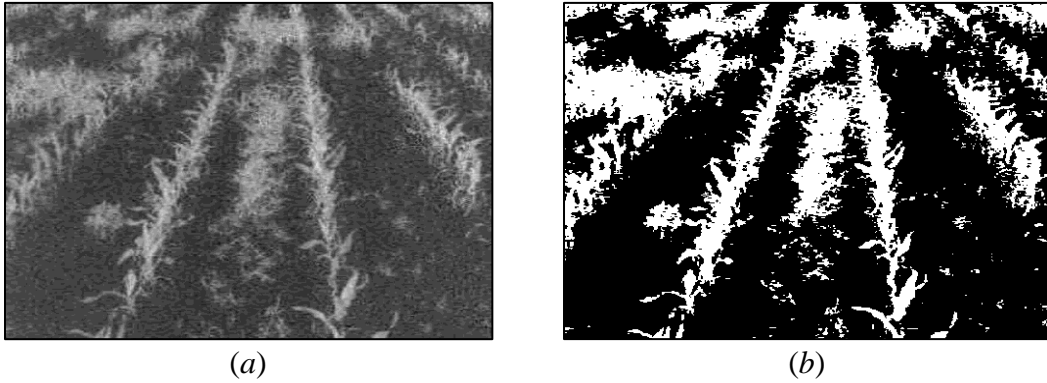


Figura 4.19 (a) Índice de vegetación con la estrategia combinada (COM); (b) Imagen binaria mediante Otsu que contiene las plantas enmascaradas y sin enmascarar.

Las figuras 4.20 (a) y (b) muestran las imágenes finales segmentadas a partir de las imágenes de la figura 4.16 (b) y (c) respectivamente, donde las etiquetas verdes y rojas insertadas sobre la imagen original corresponden a plantas sin enmascarar y enmascaradas respectivamente.

Con el fin de comprobar la validez y robustez de la estrategia propuesta de una manera más exhaustiva en el capítulo seis sección 6.3.5 se presenta los resultados obtenidos al aplicar la estrategia propuesta a un amplio conjunto de imágenes de cultivos de maíz.

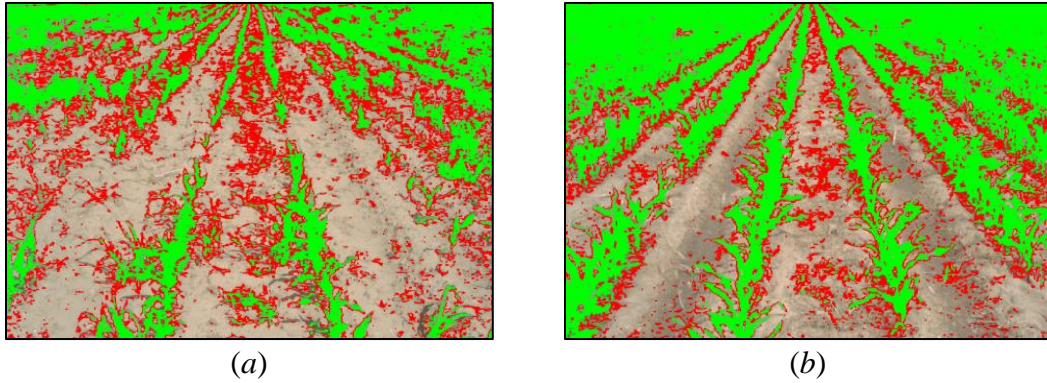


Figura 4.20 Plantas sin enmascarar (marcas verdes) y enmascaradas (marcas rojas). (a) Imagen resultado a partir de la figura 4.16(b). (b) Imagen resultado a partir de la figura 4.16(c).

De la misma manera, en Montalvo y col. (2013a) se propone una estrategia enfocada al mismo propósito de identificación de plantas cuya componente espectral verde está enmascarada por la componente espectral roja. Este enfoque se basa en la identificación de tres conjuntos mediante dos umbralizaciones. En la primera, se realiza una identificación de los píxeles que pertenecen a plantas sin enmascarar, dejando en otro conjunto aquellos píxeles que forman parte de plantas enmascaradas y del suelo. En la segunda umbralización, se divide el conjunto del suelo, que está formado por los píxeles de plantas enmascaradas y del suelo, obteniendo así los tres conjuntos deseados. Finalmente, se aplican las operaciones morfológicas de apertura (*opening*), mayoría (*majority*) y limpieza (*cleaning*) para eliminar píxeles aislados y rellenar pequeñas zonas.

5

Sistema experto: detección de líneas de cultivo y malas hierbas, guiado y solapamiento

5.1 Introducción

Tal y como se expone en el capítulo dos, el sistema de visión se compone de una cámara con sensor CCD junto con su sistema óptico asociado, algunos de cuyos parámetros intrínsecos resultan conocidos, concretamente: la longitud focal, el centro del plano de la imagen, el tamaño del píxel en micras y la resolución horizontal-vertical del CCD. En todos los casos se trata de valores nominales proporcionados por los respectivos fabricantes en sus especificaciones. Por otro lado, la cámara está situada en la zona frontal del tractor, inclinada con un ángulo de cabeceo (*pitch*, α) y a una altura determinada con respecto al suelo. También se conocen los ángulos de guiñada (*yaw*, β) y alabeo (*roll*, θ), por lo que es posible determinar las matrices de rotación y translación que nos permiten obtener los parámetros extrínsecos del sistema. Estos parámetros se determinan en función de un sistema de referencia base con respecto al cual se establecen las relaciones. En la figura 5.1 este sistema se identifica con OXYZ, mientras que el sistema de coordenadas asociado a la cámara se ha denotado como *xyz*, con el plano *xy* coincidente con el plano imagen, esto es, la propia superficie del CCD. Aunque en la figura 5.1 el sistema OXYZ se encuentra desplazado en un lateral para mayor claridad, el origen de este sistema se sitúa exactamente en el mismo plano vertical de la cámara. Mediante esta disposición geométrica, junto con los parámetros

intrínsecos, las áreas del campo se pueden identificar en el plano de la imagen por su proyección, lo que significa que dado un elemento en el campo, con su localización espacial, es posible determinar su posición relativa en la imagen. En particular, las líneas de cultivo se pueden localizar en el campo con respecto al sistema de referencia OXYZ, de suerte que mediante la correspondiente transformación entre sistemas de coordenadas, junto con el conocimiento de los parámetros intrínsecos y extrínsecos, las líneas se proyectan sobre el plano de la imagen.

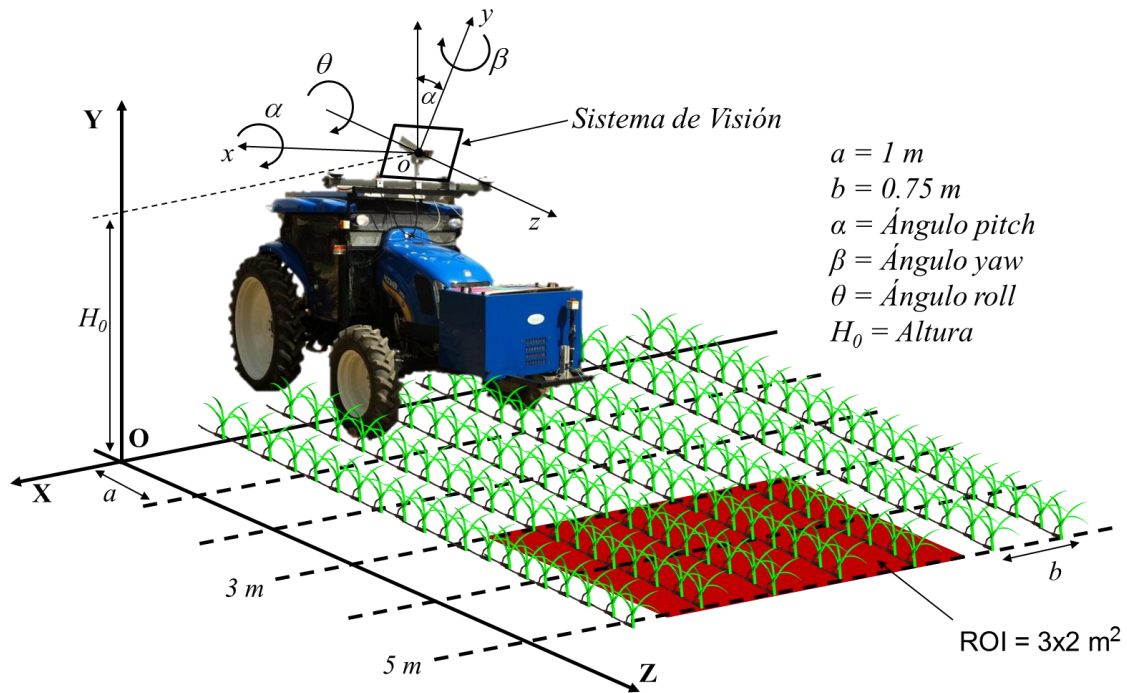


Figura 5.1 Sistema de referencia: parámetros extrínsecos.

El vehículo navega sobre un terreno real el cual presenta irregularidades. Esto provoca vibraciones y un balanceo principalmente acentuado en los ángulos de cabeceo y alabeo. El ángulo de guiñada se asume que es correcto, porque de lo contrario el tractor estaría navegando erróneamente fuera de las líneas de cultivo. Por otro lado, el espaciado entre las líneas de cultivo en el campo también se conoce. Debido a lo anterior, a menudo ocurre que la proyección hipotética, en base a la geometría conocida de las líneas de cultivo sobre la imagen, no coincide con las proyecciones reales. Esto hace que no sea posible determinar, con la precisión requerida, la proyección de las líneas de cultivo reales en la imagen y por tanto complique su identificación, lo cual impide además la aplicación correcta de tratamientos específicos por zonas.

Además, tal y como se ha expuesto en capítulos previos, la discriminación entre plantas de cultivo y malas hierbas en la imagen resulta ser una tarea difícil, ya que en ambos casos las componentes espectrales roja, verde y azul en el modelo de color RGB presentan valores similares en el momento del tratamiento. Esto significa que no es posible realizar una discriminación entre cultivo y mala hierba basándose exclusivamente en sus componentes espectrales.

Por lo tanto, la mejor opción para localizar las líneas de cultivo en la imagen de forma precisa es aprovechar la proyección geométrica de las mismas, para posteriormente realizar un ajuste fino aproximándolas a las reales mediante la extracción de verdes con índices de vegetación, de manera que las líneas proyectadas geométricamente se reubiquen sobre las nubes de píxeles verdes próximas que supuestamente corresponden a las proyecciones reales de las líneas de cultivo. De esta forma, si las líneas de cultivo están bien ubicadas, se puede identificar con precisión los píxeles a lo largo y alrededor de las mismas, mientras el resto de píxeles, que queden fuera de ellas y relativamente alejados, se pueden considerar como malas hierbas.

Este proceso de detección de líneas y malas hierbas posee un valor añadido en la aplicación desarrollada con fines agrícolas y más concretamente en AP. Por un lado permite su aplicación asumiendo las imprecisiones en la proyección de las líneas de cultivo, lo cual resulta de gran utilidad en este tipo de entornos donde los vaivenes, ligeros errores en el guiado, imprecisiones en la siembra u otros factores hacen que aparezca mucha incertidumbre en los datos. Por otro lado, como consecuencia de tales imprecisiones, no se requiere la calibración del sistema de visión (Pajares y Cruz, 2007), lo que implica que los valores para los parámetros intrínsecos son los nominales proporcionados por los fabricantes de los respectivos sistemas, por ejemplo el valor nominal de la longitud focal del sistema óptico. Respecto de los parámetros extrínsecos, sus valores se fijan de forma aproximada sin requerir una gran precisión. Concretamente, se trata de la altura H_0 en la figura 5.1 cuyo error admisible puede llegar a ser de $\pm 3\%$ y los ángulos α , β y θ que admiten desviaciones de $\pm 2^\circ$. Haber desarrollado los sistemas aptos para trabajar con este rango de errores tiene la ventaja de que la transferencia del sistema desde prototipo a producto comercial se convierte en sencilla.

Para lograr este objetivo, en la sección 5.2 se propone un sistema experto automático, el cual aprovecha el conocimiento humano, con dos módulos principales basados en técnicas de procesamiento de imágenes. El módulo de segmentación de la imagen se detalla en la sección 5.3, mientras que en la sección 5.4 se presenta el modelo geométrico necesario para determinar la posición de los objetos en el mundo real y su posición relativa en la imagen, así como un nuevo método de identificación de líneas de cultivo basado en el modelo geométrico y su corrección mediante el estimador Theil-Sen. En la sección 5.5 se explica un método de identificación de densidades de malas hierbas mientras que en la sección 5.6 se tratan los problemas relacionados con el guiado preciso del tractor, utilizando la línea de cultivo como elemento de referencia lo que en consecuencia requiere la identificación precisa de las líneas reales. Finalmente, como durante el tratamiento el tractor navega a lo largo del campo y cada imagen sirve para tratar una ROI, figura 5.1, es necesario que estas ROIs sean consecutivas en la secuencia de imágenes para evitar que se queden zonas sin inspeccionar, y en consecuencia sin tratar, o que se produzca un fuerte solapamiento entre zonas, lo que puede conducir al tratamiento reiterado de zonas. Esta problemática se aborda en la sección 5.7.

5.2 Arquitectura del sistema experto automático

La arquitectura del sistema está inspirada en el conocimiento del experto sobre la aplicación específica y también considerando los requisitos que el sistema debe cumplir. Astrand (2008) y Slaughter y col. (2008) proponen una lista de requerimientos para un sistema de guiado que pueden ser extrapolados para la detección de líneas de cultivo, ya que en esencia se trata de un problema similar. Para construir el sistema experto automático propuesto para la detección precisa de líneas de cultivo en imágenes, el conocimiento del problema se puede resumir en el siguiente conjunto de puntos:

- a) Tanto las líneas de cultivo como las malas hierbas presentan, en la mayoría de los casos, unas componentes espectrales de color similares. Además, durante el tratamiento los niveles de crecimiento del cultivo y la infestación son también similares, es decir, la altura de las plantas existentes en el campo de cultivo no es una característica que garantice una discriminación precisa.

- b) Las líneas de cultivo son una acumulación de plantas verdes siguiendo un alineamiento específico orientado hacia el punto de fuga de la imagen, propio de los sistemas basados en proyección de perspectiva (Pajares y Cruz, 2007). El cultivo se siembra con maquinaria específica que garantiza un espaciado interlineal del campo constante y conocido.
- c) Las malas hierbas, perjudiciales por ser capaces de competir con el cultivo, aparecen en rodales con una distribución irregular en el cultivo.
- d) El sistema de geometría de la cámara es conocido, es decir, se conocen los parámetros intrínsecos y extrínsecos del sistema con los errores asumibles descritos en la sección anterior.
- e) El tractor navega en terrenos que pueden presentar muchas irregularidades.

Por otra parte los requerimientos del sistema se pueden resumir como:

- f) Las líneas de cultivo deben ser identificadas con la mayor precisión posible, independientemente de la distribución de rodales de mala hierba a lo largo del cultivo. También hay que considerar que con cierta frecuencia aparecen zonas en la línea de siembra en las que se han perdido plantas debido a distintos problemas, entre ellos los errores en la siembra.
- g) El sistema de detección debe trabajar en tiempo real. Esto representa un compromiso entre la velocidad a la que navega el tractor y el coste computacional exigido a los métodos de tratamiento de imágenes. La velocidad a la que navega el tractor la fija las necesidades del tratamiento. Por ejemplo en la aplicación de herbicida la velocidad del tractor debe mantenerse en el rango de 6 a 8 km/h si se quiere que el tratamiento sea efectivo, ya que una velocidad muy alta produciría una pulverización muy fina que no llegaría a la planta mientras que una velocidad muy baja hará que la gota de herbicida sea muy gorda resbalando debido al peso por la hoja y no llegando a penetrar en la misma.

Teniendo en cuenta los conocimientos y requerimientos anteriores, y también considerando las ventajas e inconvenientes de diferentes métodos de detección de líneas de cultivo, el sistema experto automático se diseña con dos módulos principales: 1) segmentación de la imagen y 2) estimación de las líneas de cultivo. La figura 5.2

muestra el esquema de esos dos módulos con los procesos asociados correspondientes, lo que conduce a un sistema experto robusto a partir del cual se pueden detectar las densidades de malas hierbas, a la vez que realizar un proceso de guiado, considerando la continuidad de las ROIs para conseguir una cobertura total del cultivo con un tratamiento por zonas lo más preciso posible.

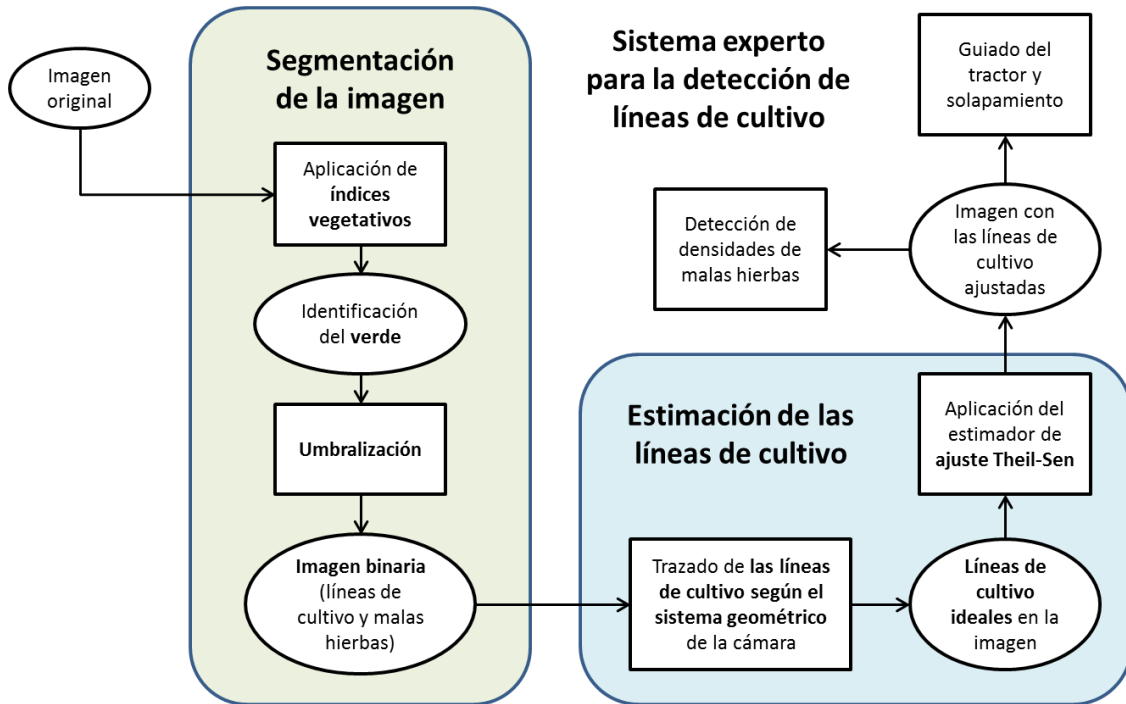


Figura 5.2 Arquitectura del sistema experto automático.

La arquitectura propuesta presenta un alto grado de flexibilidad de modo que cualquiera de los métodos puede ser sustituido por otro y se puede añadir nuevos métodos o eliminar los ya existentes. Esta flexibilidad está en línea con las propuestas de diseño descritas en capítulos anteriores.

5.3 Segmentación de la imagen

La segmentación de la imagen se centra en la separación de plantas verdes (cultivo y mala hierba) del resto (suelo, piedras y otros). De acuerdo al punto a) de la lista de conocimientos previa, y teniendo en cuenta lo expuesto en los capítulos dos y tres, secciones 2.4 y 3.3 respectivamente, la mejor opción para identificar malas hierbas y cultivo es la aplicación de índices de vegetación. En este caso y como complemento a los métodos descritos en el capítulo cuatro, se propone a continuación una combinación de métodos junto a un realzado del verde:

a) Combinación de los índices de vegetación

Por las razones expuestas en el capítulo tres sección 3.3, el índice de vegetación utilizado es el de la ecuación (5.1), basado en la combinación de índices propuesta por Guijarro y col (2011).

$$COM = w_{ExG}ExG + w_{ExGR}ExGR + w_{CIVE}CIVE + w_{VEG}VEG \quad (5.1)$$

b) Refuerzo del verde

En Romeo y col. (2012a) se propone una estrategia basada en agrupamiento borroso (*fuzzy clustering*), donde se analiza la clase que contiene los píxeles pertenecientes a las plantas verdes. Las clases contienen píxeles en el modelo RGB con tres componentes espectrales como características. Obviamente, y como es esperable, la componente espectral verde predomina sobre las otras dos. En promedio, una vez obtenido el centro de la clase para las plantas verdes, dicha componente representa el 36% con respecto a las otras dos componentes. Con este dato, y considerando que los píxeles pertenecientes a plantas deberían tener la componente verde como predominante, se incrementa el índice de verdor en COM multiplicándolo por el valor g de la ecuación (3.7), es decir, se obtiene un nuevo índice de verde de la siguiente forma: $GA = COM * g$. La multiplicación se lleva a cabo píxel a píxel, y GA se escala linealmente en el rango $[0, 1]$. Dado que g representa el porcentaje de la componente verde, la imagen resultante acentúa las partes verdes.

c) Umbralización

Dada la imagen GA transformada, el siguiente paso es su binarización para un procesamiento posterior. En la presente investigación se ha elegido el método de Otsu por su conocida eficiencia, estudiada en Meyer y Camargo-Neto (2008), con base también en el estudio de Sezgin y Sankur (2004) y explicado en el capítulo tres sección 3.4.2.

La figura 5.3(a) muestra una imagen original de un campo de maíz en el espacio de color RGB. La transformación del espacio de color utilizando el índice GA se muestra en la imagen de la figura 5.3(b). Finalmente la figura 5.3(c) muestra la imagen resultante después de aplicar el método de Otsu sobre la figura 5.3(b).

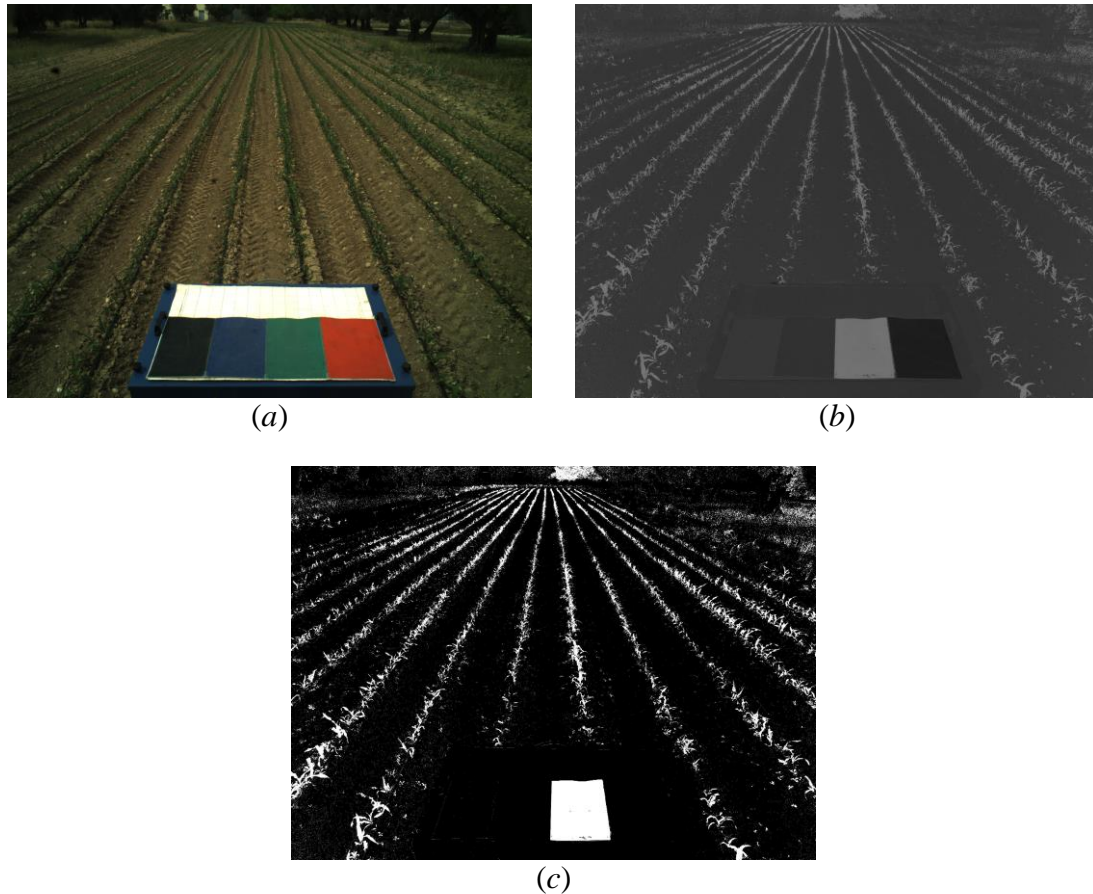


Figura 5.3 (a) Imagen original. (b) Índice GA extraído de la imagen (a). (c) Imagen binaria tras aplicar la umbralización de Otsu.

5.4 Estimación de las líneas de cultivo

Esta sección está destinada a aplicar el conocimiento de los puntos b) - e) de la sección 5.2 a la vez que se proporcionan soluciones específicas a los requerimientos f) y g) siguiendo los puntos descritos en el capítulo tres sección 3.5.

Como se ha mencionado, la navegación del tractor se realiza sobre terrenos irregulares. Por otro lado, siguiendo la estrategia combinada expuesta en la sección 5.1, para la detección de las líneas de cultivo reales se propone un ajuste basado en el estimador de regresión estadístico Theil-Sen, descrito en el capítulo tres sección 3.5.2, cuya efectividad ha sido suficientemente probada en estadística. Para realizar el ajuste se parte de las líneas estimadas mediante geometría para posteriormente aproximar estas líneas a las hileras de cultivo detectadas. A continuación se describen los pasos que completan el proceso.

5.4.1 Trazado de las líneas de cultivo estimadas mediante geometría

Como se ha mencionado previamente, dado que la ubicación de las líneas de cultivo en el campo y los parámetros extrínsecos e intrínsecos del sistema de la cámara son conocidos, es posible estimar la posición de las líneas de cultivo en la imagen obteniendo la correspondiente ecuación de la recta asociada a cada línea de cultivo (Montalvo y col., 2011; Fu y col., 1988; Hartley y Zisserman, 2006; Gottschalk y col., 2010).

Para llevar a cabo esta tarea, simplemente se necesita realizar una transformación entre sistemas de coordenadas, considerando una transformación de perspectiva en la formación de la imagen (Fu y col., 1988). Con referencia a la figura 5.1, el sistema base es el $OXYZ$ situado en el plano vertical de la cámara. De forma que un punto P en el espacio (suelo en el campo) posee coordenadas (X,Y,Z) en dicho sistema. El origen O del sistema $OXYZ$ se traslada hasta el origen del sistema $oxyz$ mediante la correspondiente matriz de traslación cuyos desplazamientos vienen dados por (X_0, Y_0, Z_0) . Realizada esta traslación, se aplican las correspondientes rotaciones elementales con los ángulos α sobre X , β sobre Y y θ sobre Z hasta hacer coincidir los ejes del sistema $OXYZ$ con los del sistema $oxyz$. Finalmente se aplica la proyección de perspectiva, considerando la longitud focal (f). De esta forma se consigue la proyección de P sobre la imagen con coordenadas (x, y) en el sistema $oxyz$ y por tanto sobre el plano de la imagen. La transformación viene dada por las ecuaciones (5.2), donde x e y son las coordenadas del punto P en el plano de la imagen expresadas en unidades de longitud.

$$x = f \left(\frac{(X - X_0) \cos \beta \cos \theta + (Y - Y_0) \cos \beta \sin \theta \cos \alpha + \dots}{(X - X_0) \sin \beta \cos \theta + (Y - Y_0) \sin \beta \sin \theta \cos \alpha + \dots} \right. \\ \left. \frac{\dots (Z - Z_0) \cos \beta \sin \theta \sin \alpha + (Y - Y_0) \sin \beta \sin \alpha - (Z - Z_0) \sin \beta \cos \alpha}{\dots (Z - Z_0) \sin \beta \sin \theta \sin \alpha - (Y - Y_0) \cos \beta \sin \alpha + (Z - Z_0) \cos \beta \cos \alpha} \right) \quad (5.2)$$

$$y = -f \left(\frac{(X - X_0) \sin \theta - (Y - Y_0) \cos \theta \cos \alpha - \dots}{(X - X_0) \sin \beta \cos \theta + (Y - Y_0) \sin \beta \sin \theta \cos \alpha + \dots} \right. \\ \left. \frac{\dots (Z - Z_0) \cos \theta \sin \alpha}{\dots (Z - Z_0) \sin \beta \sin \theta \sin \alpha - (Y - Y_0) \cos \beta \sin \alpha + (Z - Z_0) \cos \beta \cos \alpha} \right)$$

En la práctica, la traslación entre los dos sistemas de coordenadas resulta ser $(X_0, Y_0, Z_0) = (0, 0, H_0)$, siendo H_0 la altura a la que se encuentra situada la cámara con respecto al suelo, figura 5.1.

Por otro lado, si se conocen dos puntos P_1 y P_2 situados sobre una línea de cultivo en la ROI del campo, sus respectivas proyecciones, sobre el plano de la imagen, definen la recta en la propia imagen. Teóricamente, esto es posible asumiendo que el tractor está bien situado sobre el campo, además de conocer la ubicación de la ROI, sus dimensiones y las líneas de cultivo sobre ella. Se conoce igualmente la separación teórica entre las líneas en el campo, que en el caso del maíz es de 75 cm. Con ello se dispone de todos los parámetros para realizar la proyección teórica de las líneas de cultivo en la imagen.

Como se ha mencionado previamente, las coordenadas x e y en la ecuación (5.2) se expresan en unidades de longitud, por lo que desde el punto de vista de una imagen es necesario realizar la correspondiente conversión para obtener las coordenadas a nivel de píxel. Para ello, hay que tener en cuenta las especificaciones del sensor CCD que utiliza la cámara. Un sensor se caracteriza por su tamaño en términos de número de píxeles a lo largo de las dimensiones horizontal y vertical, denominadas H y V respectivamente. Cada píxel de la imagen se obtiene a partir del elemento físico que utiliza el sensor CCD tal y como se expuso en el capítulo dos sección 2.2.1. El tamaño de cada elemento es h y v expresados en unidades de longitud, normalmente en micrómetros (μm), por lo que para obtener las coordenadas en la imagen expresadas en píxeles, es necesario aplicar las ecuaciones (5.3).

$$\begin{aligned} x_p &= x / h \\ y_p &= y / v \end{aligned} \tag{5.3}$$

Nótese que los valores (X_0, Y_0, Z_0) , (X, Y, Z) , h y v han de estar expresados en la misma unidad de medida y que esta conversión ha de realizarse por cada punto en la escena o espacio 3D que se desea identificar.

La figura 5.4 muestra la proyección de las líneas de cultivo ideales tras aplicar las ecuaciones (5.2) y (5.3) con la siguiente configuración del sistema:

- Altura cámara = 2.23 m
- Número de líneas de cultivo = 4
- Distancia entre líneas de cultivo = 0.75 m
- Distancia cercana ROI = 3.3 m
- Distancia lejana ROI = 5.3 m
- Distancia focal de la lente = 10.4 mm
- Ancho CCD = 5.5 μm
- Alto CCD = 5.5 μm



Figura 5.4 Imagen original, figura 5.3(a), con la proyección de cuatro líneas de cultivo.

En resumen, tal y como se ha descrito en el párrafo anterior, es necesario conocer tanto el número de líneas como la distancia entre ellas para aplicar de forma automática un procedimiento que posicione cada línea de cultivo en el espacio tridimensional y aplique las transformaciones entre sistemas de coordenadas y de perspectiva, necesarias para obtener la proyección de las líneas de cultivo en la imagen, identificadas con sus correspondientes ecuaciones de línea recta en unidades de píxel.

5.4.2 Posicionamiento del sistema: detección de la orientación

En el apartado anterior se explica cómo posicionar las líneas de cultivo en una imagen mediante geometría bajo el supuesto de que el tractor está bien posicionado entre las líneas y las desviaciones no son significativas, cosa que en realidad raramente ocurre. Centrar el tractor de forma exacta sobre las líneas de cultivo rara vez se consigue durante la realización de una tarea agrícola debido al siguiente conjunto de factores: *a)* el terreno, con sus irregularidades, embarrado, terrones o piedras; *b)* los errores en el seguimiento de la trayectoria previamente planificada, debido a errores en las medidas del GPS cuando se capturan las medidas de cara a la planificación o cuando

se está siguiendo la trayectoria; *c*) la tracción del tractor y la resistencia de los elementos mecánicos, que en ocasiones resulta ser irregular respecto de los dos laterales del tractor, provocando desviaciones respecto de la posición ideal; *d*) errores en el sembrado, originando que las líneas de cultivo no sean exactamente rectas y que su inter-lineado no sea el teóricamente esperado de 75 cm. En Guerrero y col. (2012*c*) se presenta un estudio donde se ponen a prueba diferentes escenarios cambiando únicamente el ángulo de cabeceo (*pitch*) y manteniendo constantes el resto de parámetros del sistema (ángulos de alabeo, guiñada y altura). El estudio se realizó con campos de maíz simulados mediante la herramienta Webots (2014). El objetivo era analizar y evaluar cómo influye el posicionamiento del tractor y por ende el sistema de visión a la hora de realizar tareas agrícolas fundamentalmente cuando éstas se llevan a cabo con una flota de robots autónomos (Emmi y col., 2013). Este trabajo se amplió posteriormente en Romeo y col. (2013*b*) donde se estudió la precisión desde el punto de vista de los parámetros extrínsecos del sistema.

A continuación se proponen dos procedimientos a la hora de considerar los ángulos α , β y θ como parámetros de rotación extrínsecos relacionados con el sistema de visión.

a) Ángulos fijos: se asume que los tres ángulos son constantes, de suerte que α se obtiene mediante la correspondiente medición, mientras que β y θ se suponen cero sobre un terreno perfectamente horizontal y con el tractor correctamente orientado con respecto a las líneas de cultivo, figura 5.4.

b) Ángulos variables: En este caso se utiliza una Unidad de Medición Inercial (IMU, *Inertial Measurement Unit*). Tal y como se verá en los resultados mostrados en capítulo seis sección 6.4.1, la IMU utilizada es la MicroStrain 3DM-GX3-35 (2014), que dispone de un receptor GPS. Esta información de localización es esencial para controlar el solapamiento de las imágenes descrito en la sección 5.7. El uso de este dispositivo es apropiado para campos con las características de irregularidad antes mencionadas, ya que permite conocer las variaciones de los ángulos relacionados con los parámetros extrínsecos, controlando así la inclinación que sufre el sistema de visión respecto de su posición teórica.

Asimismo, en Conesa-Muñoz y col. (2012) se presenta una estrategia de planificación multi-camino para una flota de tractores autónomos, ya que una tarea de tratamiento de un campo real ($> 4\text{ha}$) será económicamente viable siempre que se realice con más de una plataforma autónoma de tratamiento. En esencia, cada pasada en un campo de maíz ha de realizarse siguiendo la dirección en la que se ha establecido el cultivo (*heading*). El planificador determina los puntos por lo que tiene que pasar el vehículo (plan) pero la navegación entre esos puntos tiene que estar controlada por un sistema de navegación (piloto) que puede utilizar una estrategia de movimiento entre puntos siguiendo una línea recta (Stoll, 2003) y utilizando información GPS. Cuando el cultivo es estrecho como en el caso del cereal, el vehículo puede transitar en medio del campo pisando las líneas de cultivo, ya que éste tiene capacidad de recuperarse. En este caso el planificador suministra puntos ubicados dentro del cultivo por los que tiene que pasar el vehículo y el piloto o sistema de navegación se apoya en el plan y las posiciones suministradas por el receptor GPS. Cuando el cultivo es de línea ancha como es el caso del maíz, el vehículo tiene que transitar por el medio del campo sin pisar el cultivo, puesto que éste no tiene capacidad de recuperación. En este caso el plan sólo suministra los puntos aproximados de entrada y salida al campo de modo que la navegación debe realizarse siguiendo las líneas de cultivo y manteniendo las ruedas del vehículo entre líneas para evita los daños. En definitiva, en el caso de utilizar vehículos autónomos es esencial disponer de métodos para el seguimiento de las líneas cuando el cultivo es de tipo ancho como el maíz. En el seguimiento de las líneas de cultivo, es posible que aparezcan ciertas desviaciones debido a las irregularidades del terreno, además es importante seleccionar las líneas de cultivo idóneas a seguir. Una situación de desplazamiento del tractor implica que la perspectiva cambie y por lo tanto el posicionamiento del sistema de visión también, ya que deja de estar centrado con respecto a las líneas de cultivo.

Tal y como se puede apreciar en la figura 5.5(a), el tractor no se encuentra alineado perfectamente con respecto a las líneas de cultivo por lo que está siguiendo una trayectoria desviada. Se puede ver en la figura 5.5(b) cómo en este caso aplicar una proyección geométrica utilizando un ángulo de guiñada fijo implica que las líneas estimadas no coincidan ni se aproximen a las líneas de cultivo reales.

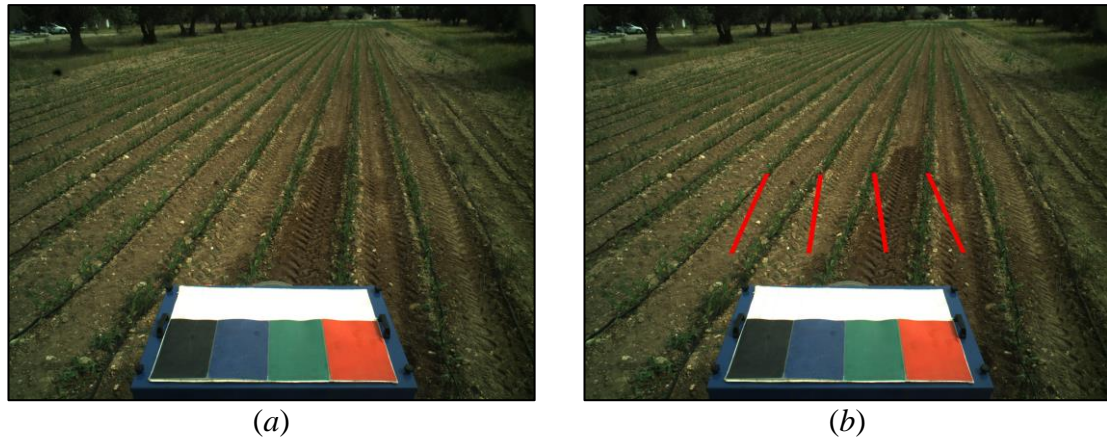


Figura 5.5 (a) Imagen original siguiendo una trayectoria desviada. (b) Proyección de líneas sobre la figura 5.5(a) calculadas en función de un ángulo de guiñada fijo.

Por ello, para solucionar esta problemática, se ha desarrollado un algoritmo de búsqueda que se encarga de encontrar el ángulo de guiñada idóneo.

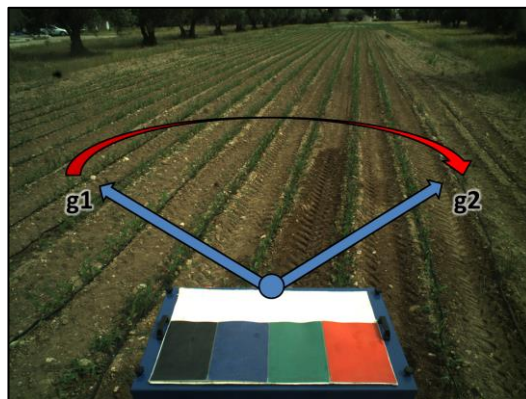


Figura 5.6 Rango de búsqueda.

La idea es sencilla, partiendo de una imagen binarizada y utilizando los valores de cabeceo y alabeo proporcionados por la IMU o fijados previamente, el algoritmo asigna valores al ángulo de guiñada barriendo un rango donde se supone que debe situarse el ángulo correcto, figura 5.6. Cuando se encuentra una coincidencia máxima de píxeles alineados, se deduce que éste es el ángulo buscado, la sinopsis del algoritmo es como sigue:

```

contador = 0
 $\forall \textit{guiñada} \in [g1, g2]$ 

    lineasEstimadas(cabeceo, alabeo, guiñada)

     $\forall \textit{pixel}(x, y) \in \textit{lineasEstimadas}$ 

         $\forall \textit{ancho} \in [x - a1, x + a2]$ 

            si  $\textit{pixel}(\textit{ancho}, y) \in \textit{Verde} \rightarrow \textit{contador} + 1$ 

angulo = guiñada  $\in \max(\textit{contador})$ 

```

- a) Se varía el ángulo de guiñada según el intervalo fijo $[g1, g2]$, dentro del cual se supone que se encuentra el ángulo buscado, y se calculan las líneas estimadas para cada ángulo de guiñada manteniendo los ángulos de cabeceo y alabeo fijos.
- b) Para cada línea estimada se obtiene el número de píxeles que forman parte del verde de la imagen (píxeles blancos en la imagen binaria) a lo largo de dicha línea y dentro de un margen de tolerancia fijo a cada lado de la línea, *a1* hacia el lado izquierdo y *a2* hacia el lado derecho. Estos márgenes permiten acotar el rango de búsqueda y mejorar el tiempo de procesamiento del método.
- c) El ángulo de guiñada óptimo es aquel que maximiza el número de píxeles verdes totales detectados en la imagen, es decir, aquel ángulo para el cual las líneas estimadas atraviesan más píxeles verdes.

Como se puede apreciar en la figura 5.7, tras la aplicación del algoritmo ante la situación mostrada en la imagen de la figura 5.5(a), las líneas estimadas se aproximan a las líneas de cultivo reales. Por lo tanto, una vez obtenidos los tres ángulos y conociendo la altura a la que está situado el sistema de visión, se pueden calcular correctamente las líneas de cultivo estimadas. Una vez hecha la corrección ya se puede aplicar el estimador de regresión de Theil-Sen para realizar el ajuste fino con el fin de proceder a la identificación de las líneas de cultivo con la máxima precisión posible.



Figura 5.7 Proyección de líneas sobre la figura 5.5(a) con un ángulo de guiñada variable.

En ambos casos, ángulos fijos o variables, los posibles errores respecto del posicionamiento de las líneas de cultivo se corrigen mediante el ajuste previsto de las líneas por el estimador de regresión de Theil-Sen. Con esta detección del ángulo de guiñada, el sistema se independiza de posibles ajustes a realizar por los agricultores como usuarios finales, haciéndolo más atractivo para una futura transferencia.

5.4.3 Corrección de las líneas estimadas: estimador Theil-Sen

Una vez calculadas las líneas de cultivo estimadas, suponiendo un sistema geométrico ideal tal y como se ha descrito en la sección 5.4.1, éstas deberían coincidir y solaparse en la imagen con las líneas de cultivo real, figura 5.4. Sin embargo, raramente hay una coincidencia perfecta o al menos válida, figura 5.5(b).

En base al estudio descrito en Guerrero y col. (2013) pueden considerarse dos tipos de relación entre las líneas estimadas y reales: *a)* coinciden; *b)* no coinciden. En el primer caso, el método de detección necesita verificar esa coincidencia. En el segundo caso, se debe aplicar una corrección sobre la línea estimada para localizar la línea de cultivo real. En cualquiera de los dos casos, bajo este enfoque, el conocimiento de la geometría del sistema, a través de los parámetros extrínsecos e intrínsecos, permite guiar al proceso de verificación y detección de líneas de cultivo.

Para verificar si las líneas estimadas coinciden o no con las líneas reales, ya que los píxeles blancos de la imagen binaria representan plantas verdes, se puede ajustar una línea recta para alineaciones específicas de píxeles, que se corresponden con alineaciones pertenecientes a líneas reales del cultivo. De este modo, dado que se tienen las ecuaciones de ambas líneas, estimada y ajustada, ya es posible verificar si existe

coincidencia entre ambas líneas. Por lo tanto, en este momento el esfuerzo se centra en los métodos de estimación de los parámetros que definen las ecuaciones de las líneas de cultivo reales, esto es la pendiente m y el término independiente b .

Un problema importante que aborda este enfoque y que constituye otra aportación significativa de la tesis, es que el método seleccionado debe hacer frente a las alineaciones específicas de píxeles además de ser lo suficientemente robusto como para evitar las desviaciones causadas por las malas hierbas que no están alineadas y se sitúan más o menos cerca de la línea de cultivo principal.

Stewart (1999) proporciona un curso orientado a la estimación de parámetros robustos en visión artificial. Dos técnicas utilizadas frecuentemente en visión son: el estimador LMS (*Least Median of Squares*) (Rousseeuw y col., 1984) y los M-estimadores (*M-estimators*) (Hampel y col., 1986; Huber, 1981). Sin embargo, es bien conocido que una gran cantidad de datos implica que las técnicas de estimación de parámetros estén fuertemente limitadas en visión artificial, incluso para las técnicas de extracción de características a bajo nivel como son la extracción de bordes. Esto a su vez implica que los problemas de estimación de parámetros en visión deberían resolverse mediante técnicas de mínimos cuadrados o, en general, mediante estimadores de máxima verosimilitud (MLE, *Maximum Likelihood Estimation*). Pero desafortunadamente los datos en visión artificial raramente se agrupan en una población estadística simple, tal y como se requiere para el uso eficaz del MLE.

Tal y como se vio en el capítulo tres sección 3.5.2, Theil (1950) propone un estimador muy simple para m usando la mediana estadística; para ello, sugiere seleccionar la mediana de las pendientes de cada posible par de puntos siempre con el mismo peso. El comportamiento asintótico de esta estimación se estudia en Sen (1968), quien extiende la definición para controlar el caso en el que los puntos tienen la misma coordenada x (rectas perpendiculares al eje x). En la definición de Sen, sólo se tienen en cuenta las pendientes de aquellos pares de puntos con distinta coordenada x , es decir, aquellas pendientes que no tienden a ∞ .

Una vez calculada la pendiente m , se puede determinar una línea a través de los puntos de muestra, asignando al término independiente b la mediana de los valores $y_i - m \cdot x_i$ (Rousseeuw y Leroy, 2003). Tal y como observó Sen, este estimador es el

valor que hace que el coeficiente de correlación por rangos del tau de Kendall (1955) llegue a ser aproximadamente cero, comparando los datos de muestra y_i con sus valores estimados $m \cdot x_i + b$.

En base a las consideraciones anteriores y a las planteadas en el capítulo tres sección 3.5.2, se selecciona el estimador Theil-Sen propuesto en Massart y col. (1997) debido a su eficiencia estadística y su robustez, incluso para imágenes con baja resolución, lo que resulta un enfoque adecuado para imágenes agrícolas que contienen líneas de cultivo.

Es bien sabido que una línea recta se representa mediante su pendiente m y su término independiente b según la ecuación (5.4).

$$Y = m \cdot X + b \quad (5.4)$$

Dada una distribución de n píxeles, el objetivo es ajustar una línea recta a la misma. Esta distribución se obtiene de la manera descrita en el capítulo tres sección 3.5 apartado b), realizando una búsqueda horizontal a lo largo de un determinado ancho para cada línea estimada, de esta forma se obtiene una distribución diferente para cada una de las líneas.

El estimador Theil-Sen evalúa pares de píxeles i y j , y calcula la pendiente sobre el conjunto de todos los posibles pares de dichos píxeles, es decir, $n(n-1)$ combinaciones posibles. Esto se lleva mediante la ecuación (5.5).

$$m = \text{med} \left(S_{ij} \mid S_{ij} = \left(\frac{y_j - y_i}{x_j - x_i + \varepsilon} \right) \right); \quad x_i \neq x_j, \quad i, j = 1, 2, \dots, n \quad (5.5)$$

donde ε tiene un valor fijo de 10^{-3} . Este parámetro se introduce con el fin de evitar las líneas exactamente verticales cuya pendiente tiende a $\pm \infty$. Aunque por definición de la ecuación (5.5) esto no puede ocurrir, cuando se implementa la ecuación algunos lenguajes de programación realizan un redondeo de los valores lo que puede provocar la aparición de líneas verticales. Otra solución para resolver este problema sería el uso de coordenadas polares.

La estimación del término independiente b se calcula como la mediana estadística de los términos independientes, que se obtienen con la pendiente m de la ecuación (5.5), que resulta ser suficientemente robusta frente a las posibles desviaciones de las líneas reales. El término independiente se calcula mediante la ecuación (5.6).

$$b = \text{med}(y_i - mx_i); \quad \forall i = 1, 2, \dots, n \quad (5.6)$$

Como se puede visualizar en las imágenes de la figura 5.8, la corrección de las líneas estimadas mediante el estimador Theil-Sen proporciona una orientación correcta, haciendo que las líneas estimadas lleguen a coincidir con las reales con la máxima precisión posible.

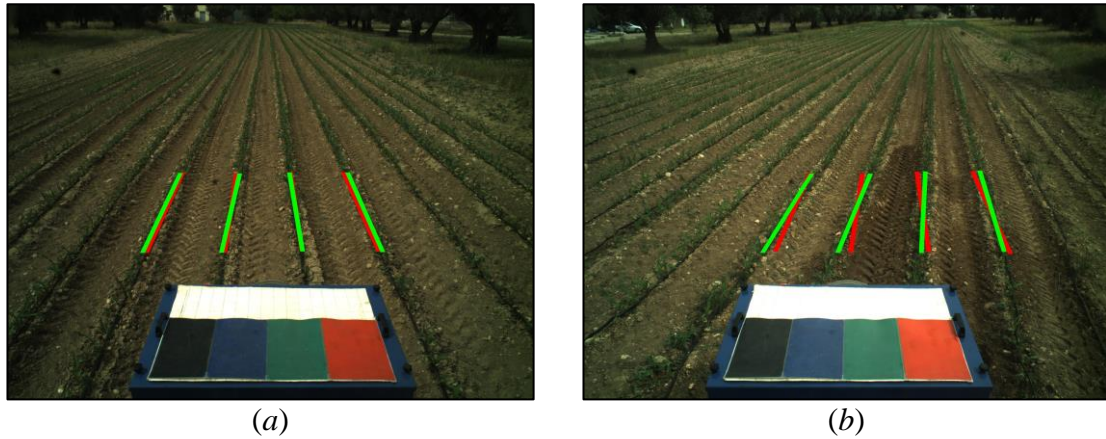


Figura 5.8 Correcciones de las líneas estimadas (líneas rojas) mediante el estimador Theil-Sen (líneas verdes). (a) Figura 5.4, siguiendo una trayectoria correcta. (b) Figura 5.7, siguiendo una trayectoria desviada.

Otro método de identificación de líneas de cultivo se basa en la regresión lineal. En Montalvo y col. (2012a) se utiliza el índice de vegetación ExG (Woebbecke, 1995) para obtener una imagen en escala de grises. Posteriormente se binariza la imagen aplicando una doble umbralización mediante el método de Otsu (1979) y con la ayuda una plantilla generada en función del conocimiento del sistema geométrico y configuración de la cámara se obtienen los píxeles que forman parte de las líneas de cultivo. La identificación de las líneas se hace calculando la recta que pasa por dichos píxeles mediante mínimos cuadrados totales (Golub, 1973; Golub y Van Loan, 1980). En el capítulo seis sección 6.4.2 se realiza un análisis comparativo de ambos métodos.

5.5 Detección de densidades de malas hierbas

Una vez las líneas de cultivo han sido detectadas, se divide la ROI real en rectángulos o celdas con la misma área. La estrategia se basa en considerar que todo píxel de verde identificado en las celdas pertenece a malas hierbas, permitiendo calcular el porcentaje de malas hierbas por celda a partir de la ecuación (5.7).

$$\text{Porcentaje} = \frac{N_{pv}}{N_{pt}} \cdot 100 \quad (5.7)$$

donde N_{pv} y N_{pt} son el número de píxeles verdes y totales en la celda respectivamente.

Todas las celdas en la ROI de la escena poseen la misma superficie, no ocurriendo lo mismo en la imagen debido a la proyección de perspectiva, siendo necesario cuadricular la ROI sobre la imagen mediante líneas verticales y horizontales. Así, para realizar la división vertical de la zona, se obtienen las bisectrices de cada par de líneas de cultivo corregidas a excepción de las dos más exteriores, para las cuales es necesario hacer uso nuevamente de la geometría del sistema de visión con el fin de mantener equidistantes las líneas. Por otro lado, para la división horizontal se utiliza la geometría del sistema para dividir la zona de estudio en función de un ancho fijado previamente, que se corresponde con la proyección de las líneas divisorias horizontales virtuales en la ROI de la escena. De esta forma se mantiene la perspectiva generada por la posición de la cámara, donde las celdas alejadas de la cámara presentan un tamaño menor que las situadas más próximas a la misma.

En las imágenes de la figura 5.9 se muestran los puntos de intersección de las bisectrices con las líneas horizontales, calculadas a 25 cm de distancia sobre el terreno. Como se puede apreciar, la proyección en perspectiva hace que las zonas de la ROI más alejadas de la cámara, pese a tener la misma área que las que se encuentran más cercanas, sean más pequeñas en lo que al número de píxeles se refiere. Como se ha indicado previamente, en Guerrero y col. (2012c) y Romeo y col. (2013b) se ha estudiado el problema de la resolución de la imagen respecto de la ROI, concluyéndose que cuando la ROI se sitúa en la parte inferior o superior de la imagen, ésta consigue la máxima cantidad de píxeles dentro, particularmente cuando se sitúa en el borde inferior.

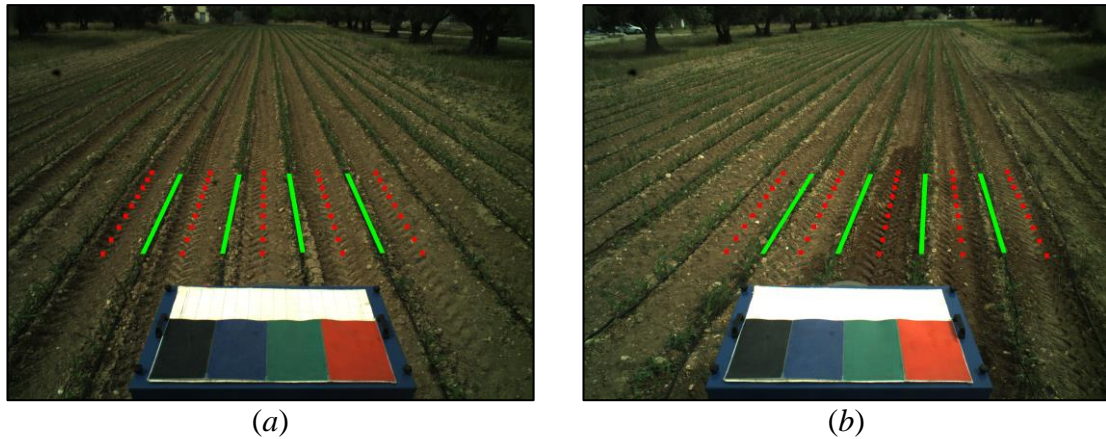


Figura 5.9 Intersecciones (puntos rojos) de las bisectrices con las líneas horizontales a 25 cm de distancia. (a) Figura 5.4, siguiendo una trayectoria correcta. (b) Figura 5.7, siguiendo una trayectoria desviada.

Una vez realizada esta división, hay que tener en cuenta el caso en el que el maíz esté en un estado de crecimiento tal que las hojas invadan los rectángulos. En este caso, como los algoritmos de segmentación identifican el verde sin distinción entre cultivo y mala hierba, es necesario descartar las dos zonas próximas a izquierda y derecha de la línea de cultivo para evitar que las hojas de cultivo de maíz que se solapan sean consideradas como malas hierbas, figura 5.10. Este ancho se fija previamente y de forma manual en función del crecimiento del maíz y el grado de precisión deseado a la hora de ajustar la línea. Por ello, a mayor crecimiento del maíz, se requiere un mayor ancho y se obtiene una precisión menor al aumentar el número de rectas candidatas a ser válidas pues existe una gran nube de puntos. En el caso contrario, a menor crecimiento de las plantas, menor es el ancho requerido y mayor precisión obtenida al acotar el número de resultados válidos. Este ancho se obtiene descartando el tanto por ciento de la distancia entre líneas con el fin de mantener la proyección en perspectiva.

En la figura 5.10 puede verse el ancho mencionado marcado mediante puntos amarillos. Para este ejemplo, el ancho se ha fijado en un 18% de la distancia entre líneas fijada a 750 mm, por lo que se descarta una zona equivalente a 135 mm a cada lado. Además, dada la proyección en perspectiva, dicho ancho es ligeramente mayor en la zona inferior que en la superior de la línea de cultivo. Como se observa, todo verde detectado en la zona delimitada por dos líneas consecutivas de puntos amarillos se corresponde con cultivo, mientras que el resto será identificado como mala hierba. Esta zona, además, delimita el ancho fijado para la obtención de la línea de cultivo corregida.

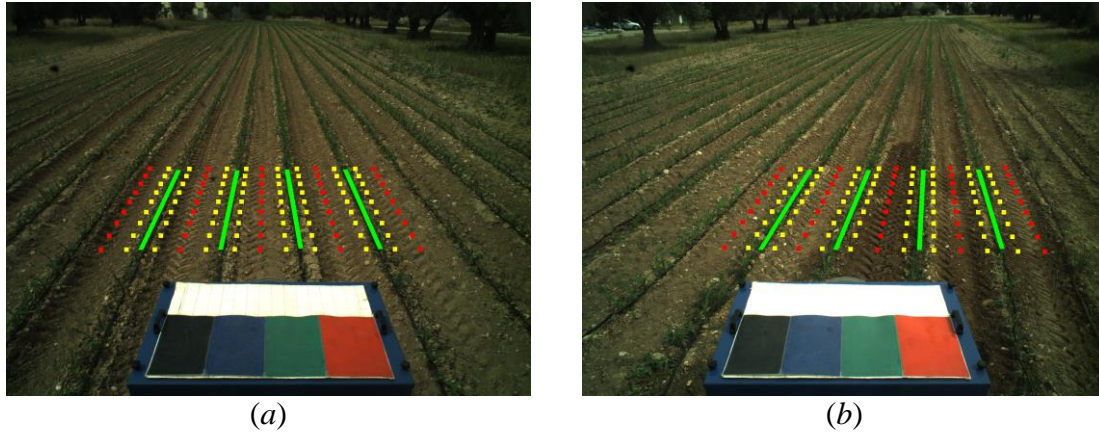


Figura 5.10 Intersecciones (puntos rojos) de las bisectrices con las líneas horizontales a 25 cm de distancia y ancho (puntos amarillos) entre los cuales se considera que los verdes pertenecen al cultivo. (a) Figura 5.4, siguiendo una trayectoria correcta. (b) Figura 5.7, siguiendo una trayectoria desviada.

Ahora es posible proporcionar para cada imagen procesada una matriz de densidades cuyas dimensiones coinciden con el número de celdas establecidas, donde cada elemento de la matriz se corresponde con el porcentaje de verdes existente en cada celda. El ancho y alto, en otras palabras el número de columnas y filas de la matriz, se determinan mediante las ecuaciones (5.8).

$$\begin{aligned}
 ancho &= 2 \cdot (\text{número de líneas de cultivo}) \\
 alto &= \left(\frac{\text{Dist. lejana ROI} - \text{Dist. cercana ROI}}{\text{Ancho cuadrículas}} \right)
 \end{aligned}
 \tag{5.8}$$

En la figura 5.11 se representa una ROI ampliada. En ella, tal y como se ha mencionado previamente, las celdas que pasan a formar parte de la matriz de densidades de malas hierbas son aquellas que no están atravesadas por las líneas de cultivo detectadas. Las celdas atravesadas son las zonas excluidas para resolver el problema del solapamiento de las hojas de maíz. Tal y como se ha especificado en el capítulo cuatro sección 4.2.2.2, la zona de estudio o ROI tiene un ancho de 3 m y una longitud de 2 m (zona inferior próxima a la cámara situada a 3.3 m y la zona superior lejana a 5.3 m) por lo que si se divide en celdas de 25 cm de largo, en este caso la matriz de densidades de malas hierbas tiene un tamaño de 8x8 resultante de aplicar las ecuaciones (5.8).

En la imagen de la figura 5.11 se han identificado las líneas de cultivo detectadas (líneas verdes), la división realizada de 25 cm (cuadrícula blanca), la zona cercana a la línea de cultivo que se descarta (zona situada entre las líneas de puntos amarillos), las rectas de separación entre dos líneas de cultivo (línea de puntos rojos), las malas hierbas detectadas (píxeles en azul claro) así como la matriz de porcentajes obtenida en este caso.

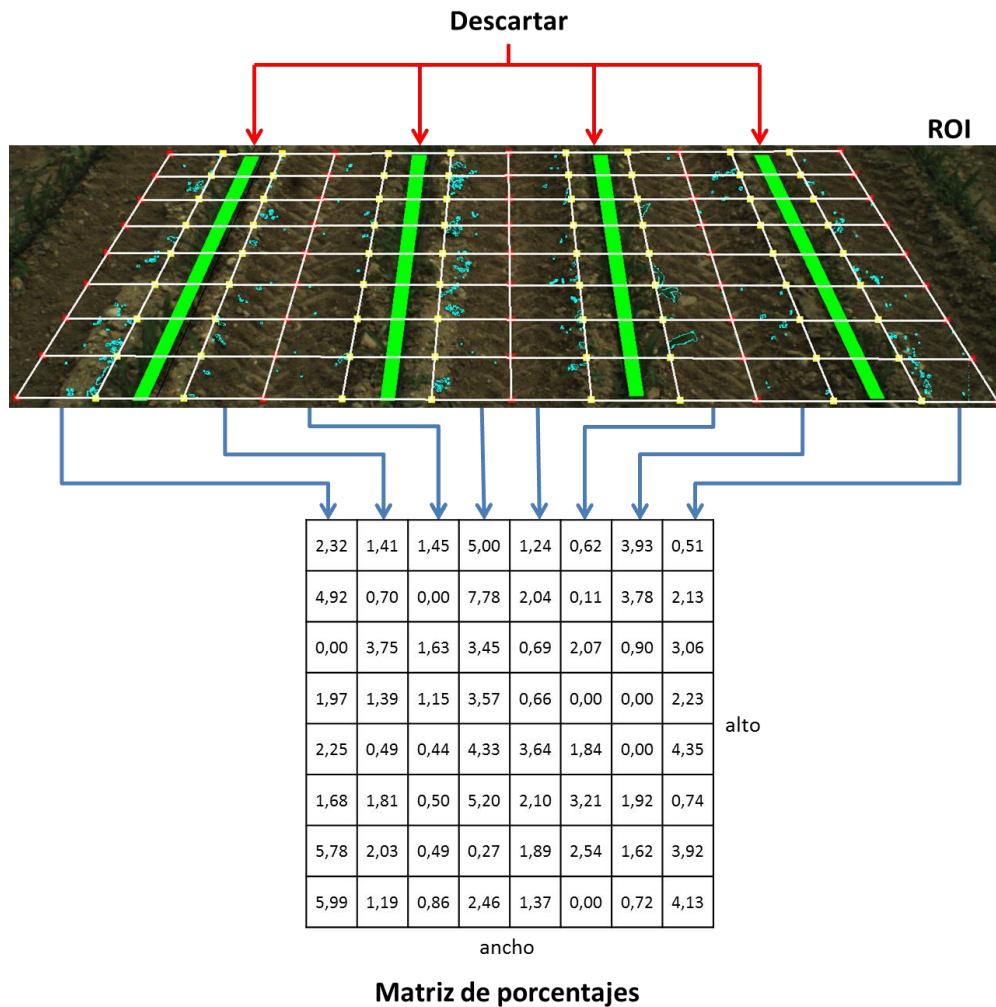


Figura 5.11 Ejemplo de la ROI con una división en cuadrículas de 25 cm de longitud.

5.6 Guiado del tractor

Suponiendo una planificación de trayectoria y su seguimiento ideal, tal y como se ha descrito en la sección 5.4.2 no habría problemas a la hora de calcular la orientación del tractor. Pero al tratarse de un tractor autónomo, cualquier fallo en la comunicación GPS o por errores en el cálculo de las posiciones puede llevar a que el tractor no siga la trayectoria correctamente.

Suponiendo un campo como el mostrado en la figura 5.12, y teniendo en cuenta las restricciones impuestas en el proyecto RHEA donde se indica que la cámara del tractor ha de abarcar un ancho de cuatro líneas de cultivo, requerimiento de la herramienta utilizada para el tratamiento de las malas hierbas, un posible plan de recorrido completo de la zona podría ser el mostrado en rojo en la figura. Como se puede observar, el plan, en este caso y suponiendo un cultivo de maíz, viene dado por los puntos de entrada y de salida del cultivo situados aproximadamente en las calles o zonas entre las dos líneas de cultivo. La trayectoria que debe seguir el tractor tiene que unir siguiendo el trazado del cultivo los puntos de entrada y salida al campo. El método propuesto tiene como finalidad la generación de esta trayectoria.

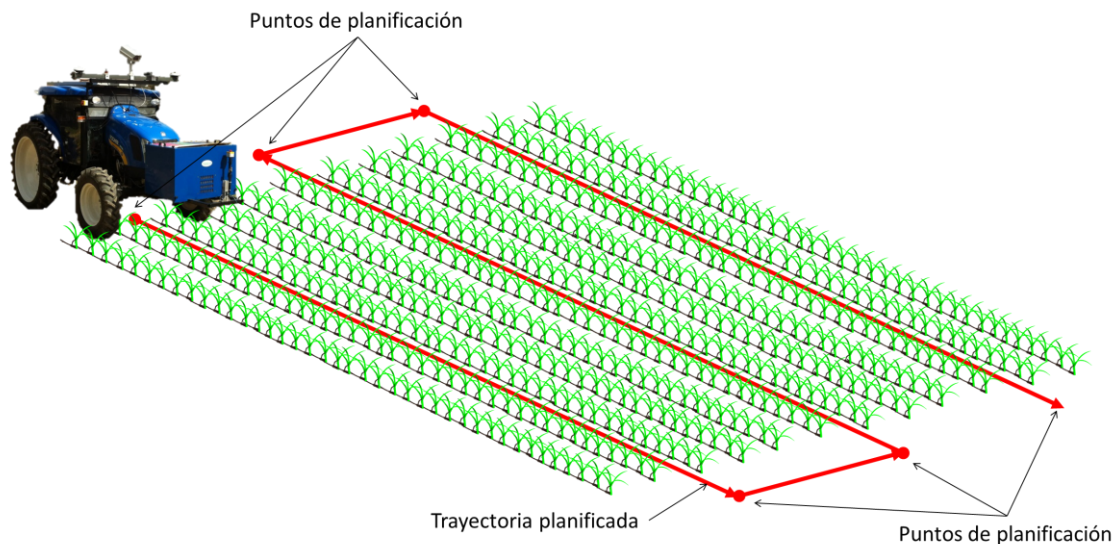


Figura 5.12 Ejemplo de planificación sobre un campo.

Como se dispone de las líneas de cultivo identificadas, tras su corrección, la bisectriz entre las dos líneas de cultivo centrales se utiliza para orientar al tractor, lo que se puede realizar de dos formas distintas dependiendo de su verticalidad. A continuación se describen ambos procedimientos:

a) Seguimiento de líneas dentro del cultivo (orientación mediante un punto): bajo la suposición de que el tractor siempre navega por el cultivo en línea recta, siempre que la bisectriz mantenga su verticalidad basta con utilizar únicamente el punto inferior de la bisectriz para determinar cuánto hay que desplazar a izquierda o derecha el tractor. De esta manera, una vez se ha desplazado la distancia necesaria, se re-ajusta la trayectoria tomando como nuevo punto de partida el calculado tras la corrección, manteniendo el mismo punto final.

Suponiendo que la bisectriz de las líneas de cultivo centrales en un modelo ideal ha de ser vertical y por lo tanto coincidir con la línea central de la imagen, se calcula el desplazamiento lateral del tractor mediante la distancia del punto inferior de la bisectriz con respecto al centro teórico de la imagen. Para ello, se obtiene el valor de la coordenada x de la bisectriz ideal, que se corresponde con el valor medio del ancho de la imagen, y se le resta el valor de la coordenada x del punto inferior de la bisectriz calculada tras la detección de las líneas de cultivo. Se utiliza este punto inferior en lugar del superior, porque tal y como muestran los estudios realizados en Guerrero y col. (2012c) y Romeo y col. (2013b), cuanto más cerca del borde inferior de la imagen se esté, mayor es la precisión obtenida.

Tras lo anterior, el resultado es un desplazamiento en píxeles, por lo tanto es necesario realizar la conversión a escala métrica. Para ello, se ha calculado la equivalencia de un píxel en milímetros mediante Webots (2014) simulando un sistema con las mismas especificaciones que el sistema de visión a bordo del tractor. Los resultados obtenidos muestran que cada píxel de la imagen situado en el borde inferior de la ROI equivale a 1.94 milímetros en la realidad, lo que da lugar a un factor de conversión de píxeles a milímetros.

La ecuación (5.9) permite calcular el desplazamiento lateral del tractor en milímetros.

$$Desplazamiento = \left(P_x(\text{bisectriz}) - \frac{\text{ancho}}{2} \right) \cdot \text{factor} \quad (5.9)$$

siendo *ancho* el valor correspondiente al tamaño horizontal de la imagen, $P_x(\text{bisectriz})$ el valor de la coordenada x del punto inferior de la bisectriz central obtenido tras la identificación de las líneas de cultivo, y *factor* es el valor de conversión de píxeles a milímetros fijado a 1.94 a partir de las simulaciones realizadas.

Una vez obtenido este valor, la interpretación del mismo indica que con un valor negativo el tractor se encuentra desplazado a la derecha y con uno positivo a la izquierda. Se observa que es necesario desplazar el tractor al lado contrario a la desviación calculada para corregir el desplazamiento observado, es decir, el tractor debe moverse en la dirección que determina la bisectriz calculada.

La figura 5.13 muestra la ROI ampliada de la figura 5.10(a). En ella se puede observar cómo la bisectriz calculada no está perfectamente alineada con respecto a la que se supone ideal, y que se corresponde con la línea vertical que atraviesa el centro de la imagen, manteniendo su verticalidad. Al situarse la bisectriz calculada a la izquierda de la ideal, el valor de desplazamiento es negativo, lo que indica que el tractor ha de realizar un ajuste a la izquierda ya que se encuentra desplazado hacia la derecha.

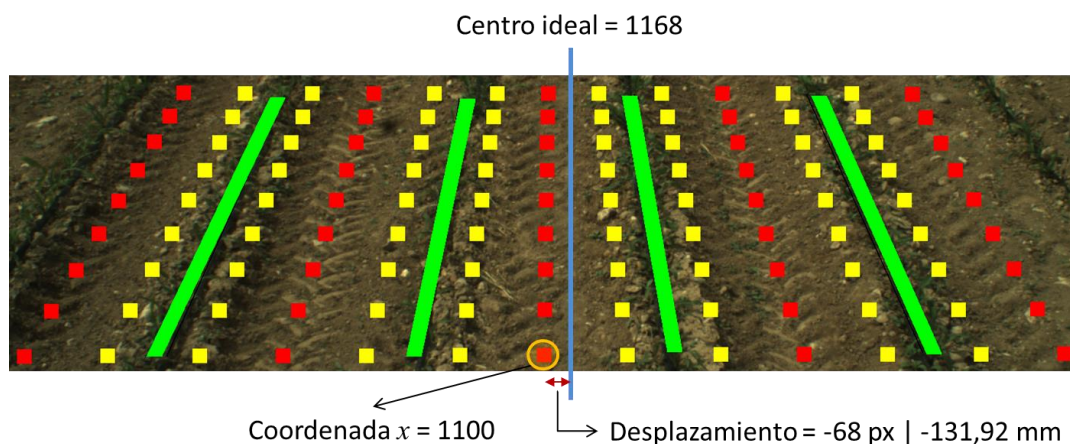


Figura 5.13 Desplazamiento calculado de la figura 5.10(a).

Como se observa en la figura 5.13, para una alineación perfecta, en este caso, sería necesario desplazar 131.92 milímetros el tractor a la izquierda. En definitiva el método propuesto es válido cuando el tractor navega en el interior del cultivo siguiendo una línea recta.

b) Aproximación al campo de cultivo (orientación mediante dos puntos): el método descrito en el punto anterior es perfectamente válido siempre que el tractor esté bien orientado en dirección al cultivo. En el caso de que el tractor no esté bien centrado y encarado para seguir una línea recta dentro del campo, la orientación mediante un único punto no es posible y hay que recurrir al uso de dos puntos. En otras palabras, cuando se trata de una aproximación del tractor hacia el campo, que requiere un giro del mismo, es necesario el uso de los dos puntos que definen la bisectriz calculada.

La figura 5.14 muestra la ROI ampliada de la figura 5.10(b). En ella se puede observar cómo la orientación de la bisectriz calculada, además de no estar perfectamente alineada respecto de la ideal, difiere considerablemente de la verticalidad deseada.

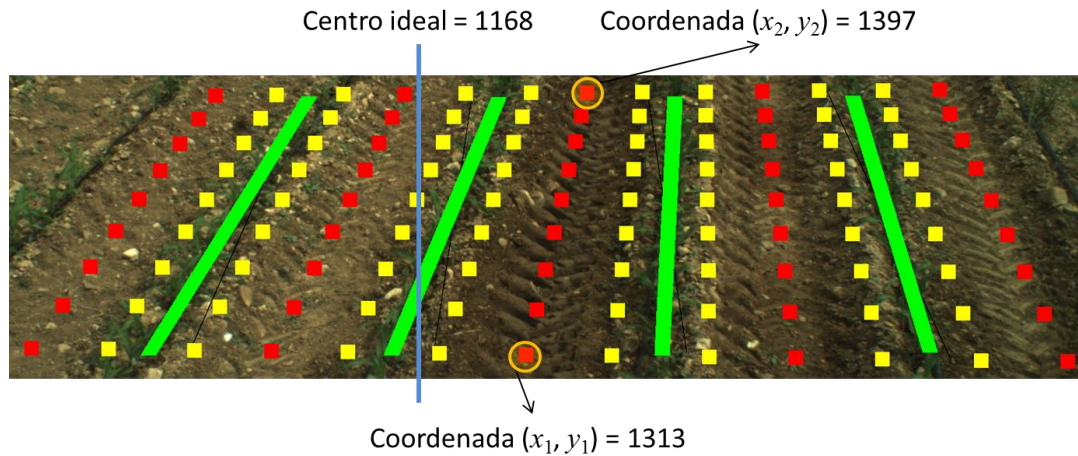


Figura 5.14 Desplazamiento calculado de la figura 5.10(b).

En este caso, en lugar de utilizar directamente los puntos suministrados por el planificador, se requiere de una corrección parcial que dirija al tractor para situarlo lo más orientado posible al cultivo, de manera que cuando el tractor esté bien centrado, de nuevo se pueda utilizar el punto final definido por el planificador que se situará justo delante.

Partiendo de que la orientación obtenida a partir de la bisectriz calculada es aquella que lleva desde el punto actual hasta el punto de salida del campo (definido por el planificador), se aplica la corrección necesaria para situar al tractor en la posición (x_1, y_1) con dirección a la coordenada (x_2, y_2) cuyas posiciones reales en el campo son conocidas y como se ha dicho anteriormente, definidas por el planificador. De esta manera se consigue una corrección parcial que permite que una vez la bisectriz calculada tenga una orientación casi vertical, se pueda re-ajustar la trayectoria desde el punto actual del tractor hacia el punto de salida, cuya posición con respecto al tractor se situará justo delante. En este instante se puede aplicar ya el método descrito en el punto anterior para guiar el tractor en la calle, es decir, desde el punto de entrada al punto de salida, ya que el tractor está orientado perfectamente hacia el cultivo.

Descritas las formas posibles de corrección, el método de guiado del tractor en un cultivo de línea ancha, tal como el de maíz, se resume como sigue:

1. Se utiliza el punto de entrada definido por el planificador para situar el tractor en la linde del campo con dirección hacia el punto de salida también establecido por el planificador.

2. Una vez situado en la posición de entrada y mientras no se haya llegado a la posición de salida, se determina a partir de la imagen obtenida si la orientación es buena o no mediante la estimación de la pendiente de la bisectriz calculada.

3. Si la pendiente indica que sí mantiene una verticalidad, se aplica el método descrito en el punto a) previo donde se calcula el desplazamiento lateral y se reajusta la trayectoria desde el punto actual con dirección al punto de salida. Se vuelve al **punto 2**.

4. Si por el contrario la orientación no es buena, se aplica el método descrito en el punto b) donde se realiza una corrección parcial posicionando el tractor en la coordenada (x_1, y_1) con dirección a la coordenada (x_2, y_2) . Se vuelve al **punto 2**.

De todo lo anterior se puede concluir que la bisectriz utilizada en este caso puede ser sustituida por la línea central calculada cuando las restricciones del campo requieran un número impar de líneas. Esta problemática ha sido también abordada en Romero y col. (2014), aunque este caso requiere más cantidad de información para el guiado, ya que utiliza una secuencia de imágenes en lugar de una única captura como en el método propuesto.

5.7 Solapamiento de imágenes

Una vez que el tractor navega por el campo siguiendo la planificación establecida por el planificador (Conesa-Muñoz, 2012) es importante para la actuación, que la matriz de porcentajes no presente solapamientos o que éste sea mínimo. Dicho de otro modo, que las ROIs en las sucesivas imágenes sean consecutivas, sin dejar huecos en el campo y sin excesivo solapamiento.

Para el control del solapamiento de las zonas de actuación se hace uso del GPS integrado en la IMU MicroStrain 3DM-GX3-35. Para ello, dado que el tractor va a seguir una trayectoria rectilínea tal y como se ha descrito anteriormente, se calcula la distancia entre la última imagen procesada y la posición actual del tractor según la ecuación (5.10).

$$Distancia = \sqrt{(Tractor_x - Imagen_x)^2 + (Tractor_y - Imagen_y)^2} \quad (5.10)$$

siendo $Imagen_x$ e $Imagen_y$ las coordenadas UTM correspondientes a la última captura de imagen realizada, y $Tractor_x$ y $Tractor_y$ las coordenadas UTM de la posición actual del tractor.

Cuando la distancia calculada (*Distancia*) se aproxima a la longitud fijada para la ROI, la cámara ha avanzado una distancia igual al largo de la ROI, en cuyo caso se garantiza la continuidad de las ROIs, y en el peor de los casos puede aparecer un cierto solapamiento debido a los errores de precisión del GPS o a pequeños desfases en las medidas por la sincronización. En cualquier caso, se trata de conseguir que el solapamiento en caso de existir sea mínimo. Esta aproximación se realiza teniendo en cuenta un factor de error calculado en función de los datos obtenidos por el GPS en estático. Para obtener este factor se mantiene el tractor en una posición fija, y durante un espacio de tiempo de 10 minutos y se almacenan las coordenadas GPS para dicha posición a una frecuencia de 4 Hz. Una vez obtenidas las 2400 coordenadas GPS, se calcula la desviación típica σ . Esta desviación típica es el error que comete el GPS y por consiguiente es la que se utiliza como el factor de error. En este caso concreto, la desviación típica calculada es de 3 cm.

Con todo esto, la primera captura de la imagen se realiza cuando el tractor afronta el campo desde la linde, a partir de ese momento se captura una nueva imagen cuando se cumple la condición enunciada en la expresión (5.11).

$$Distancia > (LongitudROI - \sigma) \quad (5.11)$$

De este modo se realiza una captura de imagen siempre que se avance la distancia asignada a la ROI teniendo en cuenta el error cometido por el GPS. Así se consigue que el solapamiento sea mínimo o inexistente, ya que es una tarea importante siempre que se quiere realizar un trabajo específico por zonas cuyo objetivo es no malgastar los recursos disponibles.

6

Análisis de resultados

6.1 Introducción

Una vez presentadas las técnicas fundamentales relativas a la segmentación de imágenes, identificación de líneas de cultivo y malas hierbas, guiado y solapamiento de las zonas de tratamiento específico, es indispensable el análisis de los resultados obtenidos con dichos métodos, tanto desde el punto de vista cualitativo como cuantitativo. Así pues, en este capítulo se presentan los resultados relativos a los diversos experimentos realizados con los distintos componentes involucrados en el proceso global, que conforman en su conjunto el sistema de visión operativo dentro del prototipo desarrollado bajo el proyecto RHEA (2014), así como parte de los desarrollos en el proyecto AGL2011. La demo final del proyecto RHEA tuvo lugar el 21 de mayo de 2014 con resultados satisfactorios.

En la sección 6.2 se describen los componentes físicos del sistema de visión con los que se ha trabajado en el desarrollo de la investigación recogida en esta memoria. En las secciones 6.3 y 6.4 se presentan los resultados obtenidos por la aplicación de las diferentes estrategias propuestas tanto para la segmentación de la imagen como de la identificación de líneas de cultivo, detección de malas hierbas, guiado del tractor y solapamiento de las zonas de interés respectivamente. Finalmente en la sección 6.5 se muestra el sistema de visión integrado en el tractor prototipo del proyecto RHEA.

6.2 Características de los componentes del sistema de visión artificial

Para la realización de las pruebas relativas a la verificación de los métodos propuestos, se han utilizado varios componentes y dispositivos con diferentes configuraciones de cámaras y lentes, cuyas especificaciones se detallan a continuación. Se describen también las especificaciones relativas a la IMU utilizada para la configuración de los parámetros extrínsecos requeridos por el sistema de visión, así como las plataformas de procesamiento utilizadas.

6.2.1 Cámaras

1. Cámara HP Photosmart R817

Se trata de una cámara comercial de dimensiones reducidas que se alimenta a través de una batería recargable del tipo ion-litio. La figura 6.1 muestra las vistas frontal y posterior de este dispositivo. Cuenta con un zoom óptico $\times 5$ y digital $\times 8$ capaz de obtener imágenes de 5.1 megapíxeles a una tasa de 30 imágenes por segundo, lo cual, unido al enfoque automático, rápido y preciso que incorpora, permite la captura de buenas secuencias de video a una velocidad por encima de la necesaria para tareas agrícolas. Las imágenes pueden ser tomadas desde una resolución mínima de 640×480 hasta una resolución máxima de 2592×1935 píxeles, mientras que el vídeo puede ser grabado con una resolución que varía desde 320×240 hasta los 640×480 píxeles. El sensor CCD tiene un tamaño de 1/2.5 pulgadas, la apertura del iris varía desde f2.8 hasta f7.6 y la velocidad del disparador se puede ajustar desde los 15 hasta 1/2000 segundos. Además, permite realizar una compensación de la exposición con diferentes modos de balance de blancos.

El mayor problema que presenta tanto esta cámara como otras comerciales de características similares, es que no permiten una conexión directa a un dispositivo para realizar un tratamiento de imágenes a tiempo real, dicha conexión se realiza vía USB o mediante la extracción de la tarjeta SD, siendo necesario en ambos casos descargar la imagen deseada antes de procesarla.



Figura 6.1 HP Photosmart R817. (a) Vista frontal. (b) Vista trasera.

2. BASLER scA1400-17FC

La principal diferencia con el modelo anterior es que se trata de una cámara industrial, la cual permite el procesamiento de las imágenes directamente al ser adquiridas, sin necesidad de descargar las imágenes previamente, es decir, no es necesario almacenar las imágenes capturadas permitiendo así una importante descarga de la memoria del sistema de procesamiento. Las imágenes obtenidas poseen una resolución de 1390×1038 píxeles y la cámara es capaz de capturar fotogramas a una tasa máxima de 17 imágenes por segundo (fps). Las imágenes se pueden adquirir con distintos modelos tanto de Monocolor como Color dependiendo del objetivo final, es por ello que para el caso de imágenes de exterior en entornos agrícolas, la cámara se ha configurado para obtener imágenes a color según el modelo Mono8. La cámara cuenta con un sensor CCD Sony ICX 285, cuyo tamaño de píxel es de $6.45 \mu m$ tanto en horizontal (h) como vertical (v). El tamaño de óptica requerida es de $2/3''$ y la comunicación se lleva cabo a través del estándar IEEE 1394b (también conocido como FireWire 800 o i.Link). La alimentación de esta cámara se realiza a través del mismo cable FireWire cuyo voltaje varía en el rango de +8 a +36 VDC. El consumo de este sistema en la configuración reseñada es de 3.0 W a 12 V. El rango de temperatura, para su correcto funcionamiento, debe situarse entre los 0° y 50° centígrados. La cámara tiene un peso de 120 g y un tamaño de $73.7 \text{ mm} \times 44 \text{ mm} \times 29 \text{ mm}$ (largo \times ancho \times alto), considerando estas dimensiones sin lente ni elementos de conexión. La figura 6.2 muestra el dispositivo descrito.



Figura 6.2 BASLER scA1400-17FC.

3. SVS-VISTEK SVS4050CFLGEA

Como en el modelo anterior, se trata de una cámara industrial y por lo tanto permite el procesamiento de las imágenes directamente, es decir, en tiempo real. Este modelo incorpora un sensor CCD Kodak KAI 04050M/C con un filtro de color Bayer con patrón GR. El tamaño del CCD es de 1". La resolución es de 2336x1752 píxeles con un tamaño de píxel de $5.5\mu m$ tanto en horizontal (h) como vertical (v) y utiliza una rosca M42 para el acople de la lente. La cámara es capaz de capturar fotogramas a una tasa máxima de 16 imágenes por segundo y de controlar el tiempo de exposición, cuyo rango varía en función de si se controla de forma interna o externa, siendo sus valores $6\mu s - 60s$ y $6\mu s - \infty$ respectivamente. Dispone de un micro-controlador interno que permite diferentes modos de ajuste para el tiempo de exposición y de distintos modos de disparadores (mecanismos para indicar que se ha de realizar una captura), por ello, se puede configurar la cámara en *Modo Libre*, que permite adquirir imágenes con la máxima tasa de adquisición o en *Modo Disparador*, el cual sincroniza la captura de imágenes en relación a un evento externo. Además, el tiempo de exposición se puede controlar mediante un puerto serie o un ancho de pulso, permitiendo aumentar el tiempo para condiciones de poca iluminación. Otra de las características que tiene esta cámara es el ajuste de las ganancias, permitiendo asignar valores diferentes a cada canal (Rojo, Verde, Azul) o realizar un ajuste automático de la misma. La lente se acopla a través de una montura tipo C y la comunicación se lleva cabo a través del estándar Gigabit Ethernet (GigE), con dos puertos de conexión de esta naturaleza, con posibilidad de enviar la misma imagen simultáneamente. La alimentación se realiza a través de un conector distinto y el voltaje es de 12 VDC con un consumo de 600 mA. El rango de temperatura para su correcto funcionamiento es de -10° hasta $+45^{\circ}$ centígrados. El tamaño de la cámara es de 55 mm x 70 mm x 70 mm (largo x ancho x alto) sin ningún

tipo de accesorio adicional, tal como lente o conectores. La figura 6.3 muestra distintas vistas de la carcasa de este dispositivo, incluyendo el CCD en su parte frontal, así como los conectores en su parte posterior.

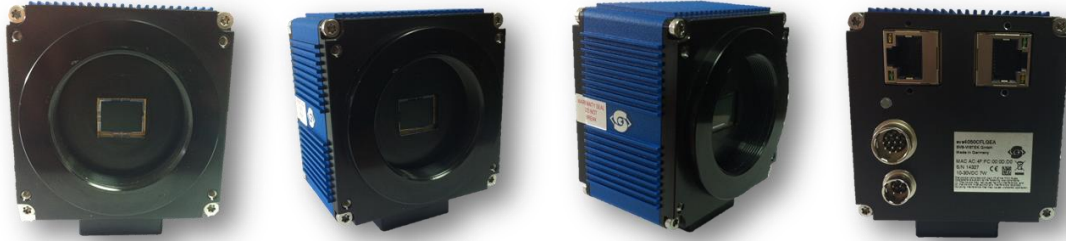


Figura 6.3 SVS-VISTEK SVS4050CFLGEA.

6.2.2 Lentes

1. PENTAX C30811KP

Esta lente es la utilizada con la cámara BASLER scA1400-17FC y por lo tanto tiene un tamaño de $2/3''$ con montura C. La longitud focal de 8.5 mm es fija y la apertura del iris se puede variar de forma manual de 1.5 hasta su completo cierre. La distancia mínima a la que ha de encontrarse un objeto es de 0.2 metros y tiene un ángulo de visión horizontal de 56.49° . El tamaño de la lente es de 42 mm x 40 mm con un peso de 120 gramos.



Figura 6.4 Lente PENTAX C30811KP.

2. SCHNEIDER KREUZNACH CINEGON 1.9/10-0901

Esta lente (Schneider Kreuznach, 2014a) junto con su correspondiente adaptador es la utilizada con la cámara SVS-VISTEK SVS4050CFLGEA de acuerdo al tamaño necesario de $1''$ del CCD, requisito que cumple dicha lente. El adaptador es necesario debido a que la lente utiliza una montura C y la cámara tiene una rosca M42. El rango

de longitud de onda que admite esta lente es el comprendido entre 400-1000 nm, lo que equivale al espectro visible y parte del infrarrojo cercano. La longitud focal es fija a 10.4 mm y la apertura del iris varia de forma manual de 1.9 a 16 mm. El peso es de 136 g con un tamaño de 50.2 mm \times 34 mm. La figura 6.5 (a) muestra la lente sin adaptadores, mientras que en (b) se muestra esta misma lente con los correspondientes adaptadores tanto para el filtro como para su acoplamiento a la carcasa de la cámara.



Figura 6.5 Schneider Kreuznach Cinegon 1.9/10-090. (a) Lente sin adaptadores. (b) Lente con adaptadores para el filtro y la cámara.

Dado que esta lente admite longitudes de onda que no pertenecen al espectro visible y debido a que la cámara es muy sensible a la radiación NIR la cual contamina los tres canales espectrales (Rojo, Verde, Azul) en especial al rojo debido a su proximidad espectral, es necesario el uso de un filtro de corte UV/IR. Como ya se mencionó en el capítulo tres sección 3.2.2, el filtro utilizado en este caso junto a un adaptador que facilita el acople con la lente es el modelo Schneider UV/IR 486 (Schneider Kreuznach, 2014b) mostrado en la figura 6.6(a), el cual corta tanto la radiación ultravioleta (UV) a 300-375 nm como la infrarroja (IR) a 750-1035 nm, permitiendo el paso de longitudes de onda que varían en el rango de los 390 nm \pm 15 nm a los 690 nm \pm 20 nm, tal y como se especifica en su curva de respuesta característica, figura 6.6(b).

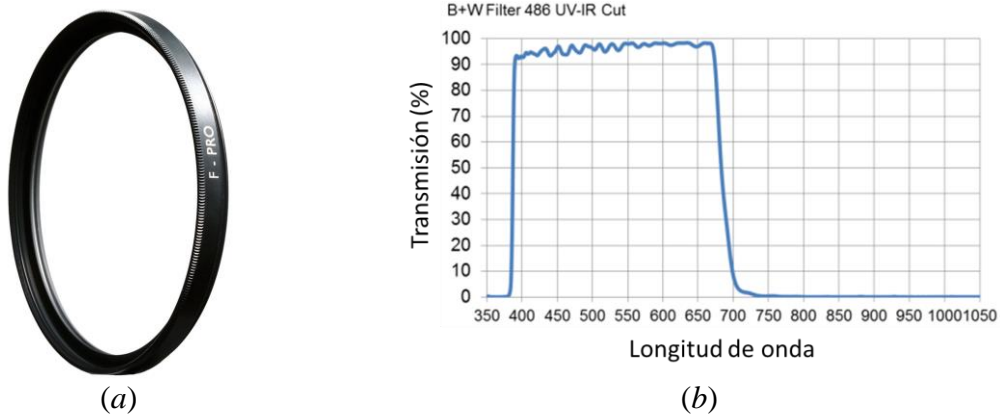


Figura 6.6 (a) Filtro Schneider UV/IR 486. (b) Curva de respuesta del filtro Schneider UV/IR 486.

6.2.3 Unidad de medición inercial

La IMU utilizada es la MicroStrain 3DM-GX3-35 (AHRS) con GPS, la cual combina un sensor de tecnología electromecánica (Microelectromechanical Systems, MEMS) y un receptor GPS integrado de alta sensibilidad. Incorpora un acelerómetro, giróscopo y magnetómetro, todos ellos triaxiales (por tanto de tres ejes), además de sensores de temperatura y un procesador integrado que ejecuta un algoritmo de fusión sofisticado que proporciona la orientación, inercia y las mediciones GPS. Este modelo ofrece una amplia gama de medidas completamente calibradas tales como la aceleración, índice angular, campo magnético, ángulos de Euler (cabeceo o *pitch*, alabeo o *roll*, guiñada o *yaw*), matriz de rotación y cuaternión. Los datos GPS incluyen la posición LLH, velocidad NED, posición y velocidad ECEF, datos DOP, hora UTC, hora GPS, información del reloj, ajuste GPS y SVI. Todas estas cantidades se compensan por temperatura y matemáticamente se alinean según un sistema de coordenadas ortogonal. Además, las frecuencias angulares se corrigen mediante sensibilidad G escalándose a un factor no lineal de tercer orden. La arquitectura de este dispositivo ha sido diseñada para eliminar errores comunes tales como la histéresis inducida por los cambios de temperatura y la sensibilidad a las variaciones del voltaje de entrada. La desviación del giróscopo se elimina con el modo AHRS mediante el cual se referencia el Norte magnético y la gravedad terrestre que son utilizados para compensar el giróscopo con un sesgo. La conexión se realiza mediante USB o RS-232, siendo esta última la elegida para el sistema propuesto. Con respecto a la precisión de los ángulos proporcionados, ésta es de $\pm 0.5^\circ$ en estático y $\pm 2^\circ$ en dinámico. El peso de este dispositivo es de 23 gramos y tiene un tamaño de 44 mm x 24 mm x 14 mm (largo x ancho x alto)

excluyendo los elementos de montaje. Considerando los soportes para su anclaje el ancho se incrementa hasta los 37 mm.



Figura 6.7 MicroStrain 3DM-GX3-35(AHRS) con GPS.

6.2.4 Plataformas de procesamiento

1. PC ASUS N61Jq

Se trata de un portátil ASUS modelo N61Jq con Windows Home Premium de 64 bits el cual cuenta con un procesador Intel® Core™ i7 720QM de 1.6 GHz con 4 GB de memoria RAM y una tarjeta gráfica modelo ATI Mobility Radeon HD 5730 con 1 G DDR3 VRAM dedicada. Además de la conexión Gigabit Ethernet para trabajar directamente con la cámara SVS-VISTEK SVS4050CFLGEA incorpora un slot Express card que permite la conexión de la cámara BASLER scA1400-17FC.

2. NI cRIO-9082 RT

Los sistemas CompactRIO tienen una arquitectura hardware robusta que consiste en un controlador embebido para comunicación y procesamiento, un chasis reconfigurable que aloja la FPGA programable por el usuario, módulos de E/S intercambiables en vivo y software gráfico LabVIEW para programación rápida en Windows, FPGA y sistemas en tiempo real. Este tipo de sistemas los distribuye la empresa National Instruments (2014).

El modelo 9082 cuenta con una CPU Intel Core i7-660UE de 1.33 GHz con dos núcleos, cuya caché de nivel 2 es de 256 KB por núcleo y en el caso de nivel 3 es de 4 MB compartida por ambos núcleos. El almacenamiento es no volátil de 32 GB y dispone de 2 GB DDR3 800 MHz de RAM. Este sistema funciona bajo Windows Embedded Standard 7 y Real-Time, este último elegido por la necesidad de trabajar en tiempo real. Debido a su robustez, el sistema puede trabajar a una temperatura que varía

de los 0 °C hasta los 55 °C, haciéndolo apropiado para tareas agrícolas cuyos sistemas pueden estar sometidos a temperaturas cambiantes y extremas. Además cuenta con un grado de protección de nivel IP20.

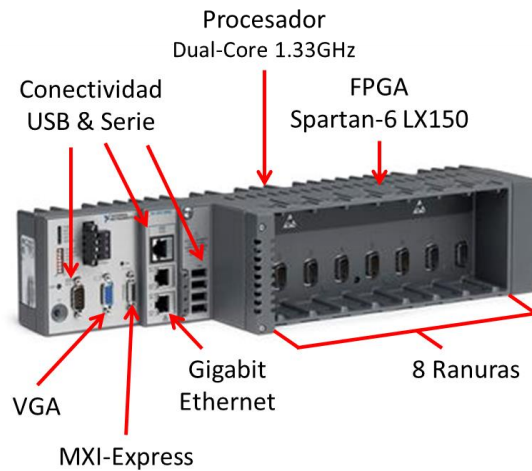


Figura 6.8 Sistema NI cRIO-9082 RT.

6.3 Análisis de la segmentación de imágenes

En esta sección se analizan los resultados proporcionados por los algoritmos relacionados con la segmentación de imágenes que han sido propuestos a lo largo del capítulo cuatro, siguiendo el mismo orden.

6.3.1 Identificación de imágenes con alta intensidad lumínica

Para este caso se ha utilizado un conjunto de 500 imágenes reales tomadas en campos de cultivo de maíz y adquiridas con la cámara BASLER scA1400-17FC durante abril/mayo de 2011 en un campo de 1.7 hectáreas de extensión situado la Finca Experimental del CSIC en La Poveda (Arganda del Rey, Madrid). La adquisición de las imágenes se dividió en distintas series de aproximadamente 100 imágenes por serie, dejando un espacio temporal de cinco/seis días entre cada serie, es decir, fueron obtenidas bajo diferentes condiciones medioambientales y de iluminación, incluyendo niveles altos y normales de saturación del CCD y con diferentes estados de crecimiento tanto de las plantas de maíz como de las malas hierbas. Las imágenes digitales fueron almacenadas como imágenes de 24-bits a color, con una resolución de 1392x1038 píxeles en el modelo de color RGB y en formato BMP. Dos imágenes representativas se muestran en la figura 4.3.

Del conjunto de 500 imágenes, se seleccionaron 250 aleatoriamente y fueron utilizadas para la fase de entrenamiento tal y como se ha descrito en el capítulo cuatro sección 4.2.1.1, y las 250 restantes fueron utilizadas para la fase de validación, es decir, para clasificarlas como pertenecientes al conjunto C1 o C2 utilizando el valor de umbral fijo a 0.6 definido en el capítulo cuatro sección 4.2.1.1.

Con el fin de validar la estrategia propuesta, cada imagen se clasifica y de ella se obtiene su porcentaje de verde (PG). Basado en esto se calculan las siguientes cantidades:

- VP (verdaderos positivos): número de imágenes clasificadas como C1 con un $PG < 12\%$.
- VN (verdaderos negativos): número de imágenes clasificadas como C2 con un $PG \geq 12\%$.
- FP (falsos positivos): número de imágenes clasificadas como C1 con un $PG \geq 12\%$.
- FN (falsos negativos): número de imágenes clasificadas como C2 con un $PG < 12\%$.

Con estas cantidades se calcula el porcentaje de clasificación correcto (PCC) tal y como se describe en la ecuación (6.1) (Sneath y Sokal, 1973).

$$PCC = \frac{VP + VN}{(VP + VN + FP + FN)} \quad (6.1)$$

Con el conjunto de las 250 imágenes para validación se han obtenido los siguientes valores: $VP = 138$; $VN = 98$; $FP = 8$; $FN = 6$. De este modo, el valor PCC obtenido es 94.4%. Este valor representa un alto porcentaje de imágenes correctamente clasificadas. Sólo 8 imágenes fueron clasificadas como FP, debido al efecto del cielo, ya que se trata de una zona más iluminada que el suelo o las plantas y por lo tanto el valor de intensidad obtenido aumenta.

Estas imágenes han sido procesadas con la Image Processing Toolbox de Matlab R2009b (TheMathworks, 2014) mediante el portátil ASUS obteniendo un tiempo promedio de procesado por imagen de 0.8 segundos. Además, con el fin de aplicar el método propuesto para su procesamiento en tiempo real, se realizó una implementación

del algoritmo en C, mediante MS Visual Studio 2008, y se compiló como una DLL que fue embebida como un módulo de LabVIEW dentro de la cRIO obteniendo un promedio de tiempo de procesamiento por imagen de 25 milisegundos. A la vista de ambos resultados, se infiere claramente la disminución respecto al tiempo de proceso en ambas implementaciones.

6.3.2 Vignetting y balance de blancos

Para el estudio del *vignetting* y del balance de blancos se realizó una selección de 200 imágenes en formato BMP capturadas con la cámara SVS-VISTEK SVS4050CFLGEA en la Finca Experimental del CSIC en La Poveda (Arganda del Rey, Madrid) durante mayo de 2013 bajo distintas condiciones de iluminación, nubosidad y densidades de maíz y malas hierbas. Estas imágenes han sido procesadas con la cRIO-9082 en tiempo real mediante una DLL implementada en C y embebida dentro de un módulo de LabVIEW 2011. Con el fin de estudiar la mejora total de la imagen, no se realizó el estudio sobre la ROI, sino que se tuvo en cuenta la totalidad de la imagen ya que, como se ha mencionado reiteradamente, es sobre todo en los bordes exteriores de las imágenes donde más se acentúa el problema derivado del efecto *vignetting*.

Con el fin de comprobar la validez y precisión del método propuesto, se seleccionan aleatoriamente el 25% de las 200 imágenes disponibles, haciendo un total de 50. Para cada una de estas 50 imágenes se construye una imagen *ground-truth* a partir de su imagen binarizada. La binarización se obtuvo tras aplicar una segmentación con el índice ExG descrito en el capítulo tres sección 3.3 seguida de una umbralización basada en el método de Otsu descrito en el capítulo tres sección 3.4.2. Esta imagen binaria se retocó manualmente, así los píxeles aislados o grupos de píxeles se volvieron a etiquetar como píxeles blancos o negros según el criterio de un experto.

Después se comparó la imagen binaria obtenida mediante ExG y Otsu (ExG&Otsu), con y sin modificar la imagen original con ambos métodos, con la correspondiente imagen *ground-truth* obtenida. En base a esto se calcularon las siguientes cantidades:

- BV (blancos verdaderos): número de píxeles clasificados como blancos en la imagen ExG&Otsu y también en la *ground-truth*.
- NV (negros verdaderos): número de píxeles clasificados como negros en la imagen ExG&Otsu y también en la *ground-truth*.
- BF (blancos falsos): número de píxeles clasificados como blancos en la imagen ExG&Otsu y negros en la *ground-truth*.
- NF (negros falsos): número de píxeles clasificados como negros en la imagen ExG&Otsu y blancos en la *ground-truth*.

Con las siguientes cantidades se calcula el porcentaje de clasificación correcto (PCC) tal y como se describe en la ecuación (6.2) (Sneath y Sokal, 1973).

$$PCC = \frac{BV + NV}{(BV + BF + NV + NF)} \quad (6.2)$$

La tabla 6.1 muestra los valores PCC medios de las 50 imágenes analizadas para la identificación del verde, junto con el tiempo de procesamiento promedio en milisegundos (ms) para la imagen original binarizada y la imagen binarizada tras ser tratada con ambos métodos.

Tabla 6.1 Valores PCC y tiempos de procesamiento medio para la identificación del verdor en imágenes originales y tratadas con el *vignetting* y el balance de blancos.

	ExG&Otsu	<i>Vignetting</i> + ExG&Otsu	Balance de blancos + ExG&Otsu
PCC (%)	80.3	87.3	81.6
Tiempo (ms)	82.5	132.2	95.7

En los resultados mostrados en la tabla 6.1 se observa que la imagen con la corrección del *vignetting* mejora sustancialmente con respecto a la extracción de verdes mientras que el balance de blancos apenas resulta efectivo. Esto se debe en gran medida a que el balance de blancos mejora la imagen de forma que es más atractiva para el ojo humano, pero no mejora los valores de cada canal espectral ya que aumenta o disminuye todos los canales proporcionalmente y no se centra en uno concreto como ocurre en la corrección del *vignetting*. Los resultados visuales de ambas técnicas son los mostrados en las figuras 4.6 y 4.7 para el *vignetting* y el balance de blancos respectivamente.

En cuanto al tiempo de procesamiento hay que destacar que el proceso de corrección del *vignetting*, aplicando las ecuaciones descritas en el capítulo cuatro sección 4.2.2.1, llega a ser un proceso relativamente costoso debido al continuo cálculo de las distancias. Dado que el tamaño de imagen es siempre constante y no varía de un instante a otro, este proceso se ha mejorado para trabajos en tiempo real añadiendo los valores de distancia en un fichero externo que es precargado por el algoritmo, obteniendo en este último caso un tiempo de proceso promedio de 94.3 ms, por lo que mejora el anterior en aproximadamente un 29%.

6.3.3 Ajuste del tiempo de exposición

La estrategia propuesta en el punto a) del capítulo cuatro sección 4.2.2.2 presenta una serie de problemas que impiden su uso en aplicaciones agrícolas. Por lo tanto, el estudio del tiempo de exposición se centra en la estrategia b) propuesta en la misma sección, que está basada en la intensidad lumínica y aparece descrita en Guerrero y col. (2014). Se presenta una comparativa de esta estrategia con una simplificación del método propuesto en Montalvo y col. (2013b).

Con el fin de calcular el rango de valores límite para cada método, se dispone de un conjunto de 1834 imágenes que se dividen según el criterio de un experto en los siguientes tres grupos, según el grado de exposición del sensor a la iluminación exterior: a) subexpuestas; b) normales; c) sobreexpuestas. Las imágenes se adquirieron con la cámara SVS-VISTEK SVS4050CFLGEA y se procesaron con la cRIO-9082 en tiempo real mediante una DLL implementada en C y embebida dentro de un módulo de LabVIEW 2011. Las imágenes pertenecen a campos de cultivo de maíz y fueron captadas en la Finca Experimental del CSIC en La Poveda (Arganda del Rey, Madrid) en mayo de 2013, obteniendo imágenes con el tractor prototipo del proyecto RHEA tanto parado como en movimiento. La ROI utilizada para llevar a cabo el estudio posee las dimensiones y características definidas en el capítulo cuatro, sección 4.2.2.2.

Los valores proporcionados en Montalvo y col. (2013b) se obtienen a través del estudio de los histogramas de las imágenes con diferentes condiciones lumínicas. Ya que en la ROI los canales predominantes son el Rojo y el Verde, sólo se estudian estos dos y no el Azul, puesto que éste último está directamente relacionado con las partes correspondientes al cielo. De acuerdo a los valores medios de cada conjunto, el nuevo

tiempo de exposición (TE) basado en el tiempo de exposición actual (TE_{actual}) se calcula como sigue:

$$\begin{aligned} & \text{mientras } (m_i > m_{ui}) \text{ o } (m_i > m_{li}) \\ & TE = \left(\frac{m_{Rref} + m_{Gref}}{m_R + m_G} \right) TE_{actual} \\ & \text{en otro caso } TE = TE_{actual} \end{aligned} \quad (6.3)$$

donde i representa los canales Rojo (R) y Verde (V) y m_i se refiere a los correspondientes valores medios (m_R, m_G) del histograma a analizar. Los subíndices u y l significan superior (*upper*) e inferior (*lower*) y son los valores límite, que se corresponden a $m_{uR} = 229.2$, $m_{uG} = 225.4$, $m_{lR} = 26.7$ y $m_{lG} = 24.6$ según los valores superiores e inferiores obtenidos de los conjuntos de imágenes subexpuestas y sobreexpuestas respectivamente. Los valores m_{Rref} y m_{Gref} se obtienen a partir de la observación de una serie de histogramas usados como referencia, seleccionados según el criterio de un experto, obteniendo finalmente los valores de 82.2 y 75.4 respectivamente.

Con respecto a la estrategia b) propuesta en el capítulo cuatro sección 4.2.2.2 y resumida en la ecuación (6.4) por simplicidad, los valores de los umbrales se determinan obteniendo los valores límite de cada conjunto para conseguir los valores medios superior e inferior necesarios para adquirir imágenes cuya intensidad lumínica sea considerada como normal según el criterio del experto mencionado.

$$\begin{aligned} & \text{si } (m_l > m_u) \rightarrow TE = TE_{actual} - (TE_{actual} / 2) \\ & \text{si } (m_l < m_l) \rightarrow TE = TE_{actual} + (TE_{actual} / 2) \\ & \text{en otro caso } TE = TE_{actual} \end{aligned} \quad (6.4)$$

donde m_u y m_l representan los límites superior (*upper*) e inferior (*lower*) fijados a 0.1 y 0.4 respectivamente. El umbral entre las imágenes subexpuestas y normales se sitúa muy próximo a 0.1, siendo el valor máximo de las imágenes subexpuestas 0.09 y el valor mínimo de las imágenes normales 0.13. En el caso del umbral entre las imágenes normales y sobreexpuestas, los valores máximo y mínimo de cada conjunto son 0.38 y 0.41 respectivamente, ambos próximos a 0.4. Con esto se verifica que los valores asignados en el capítulo cuatro sección 4.2.2.2 son válidos para las imágenes relacionadas con las aplicaciones agrícolas en cuestión.

La figura 6.9 muestra la evolución experimentada por el tiempo de exposición con esta estrategia durante la captura de 1025 imágenes. Tal y como puede observarse, el tiempo de exposición se mantiene prácticamente constante hasta que una nube ensombrece la escena, momento en el que el tiempo de exposición aumenta considerablemente con el fin de obtener una imagen menos oscura y por tanto de mayor contraste y calidad. El círculo rojo muestra una zona de variación ampliada donde puede verse cómo el método, por su diseño, necesita entre dos y tres imágenes por cada cambio brusco para ajustarse correctamente.

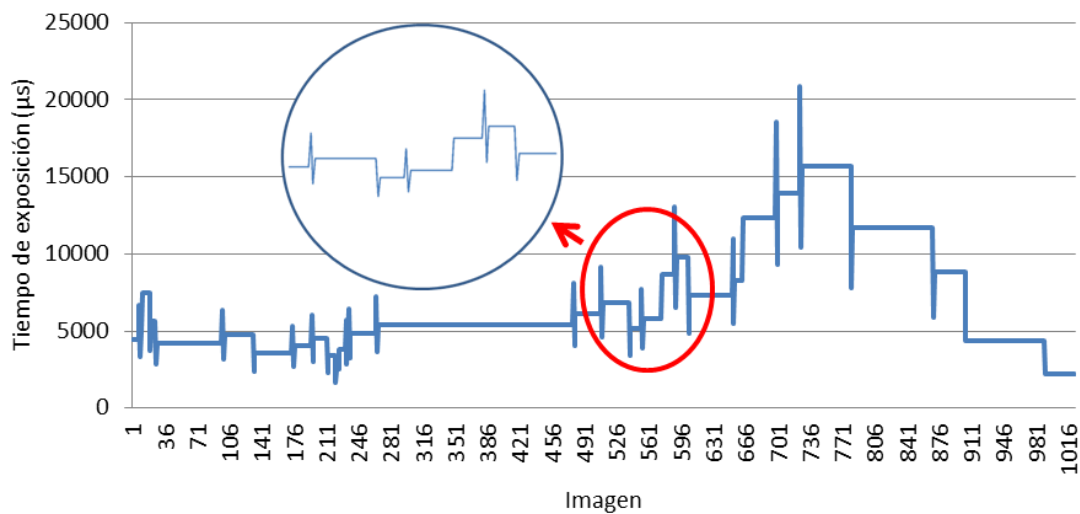


Figura 6.9 Variación del tiempo de exposición.

Una vez se han calculado los umbrales, la comparación se realiza a través de la precisión en la detección de verdes tras aplicar la técnica de las SVM, descrita en el capítulo cuatro sección 4.4, así como de la tendencia de cada método a adquirir imágenes subexpuestas, normales o sobreexpuestas.

6.3.3.1 Precisión en la detección de verdes

Con el fin de capturar imágenes bajo distintas condiciones de iluminación causadas por la variabilidad debida a diferentes densidades de nubes, se han seleccionado diez posiciones diferentes en el campo. En cada posición se han aplicado ambos métodos de ajuste, ecuaciones (6.3) y (6.4), obteniendo diez imágenes de cada uno. También para cada posición se ha forzado el tiempo de exposición con diferentes valores para capturar diez imágenes subexpuestas, diez normales y diez sobreexpuestas. Los valores del tiempo de exposición se han fijado entre 1 y 20 ms para imágenes

subexpuestas, entre 20 y 80 ms para imágenes normales y entre 80 y 100 ms para imágenes sobreexpuestas. Se adquirieron un total de 500 imágenes.

Del conjunto de imágenes anterior se comprueba la precisión en la detección del verdor en una selección de 40 imágenes. Para cada una de las 40 imágenes se construye una imagen *ground-truth* de la ROI una vez se han binarizado aplicando las SVM descritas en el capítulo cuatro sección 4.4. Como en el estudio del *vignetting* y del balance de blancos, esta imagen binaria se retoca manualmente renombrando los píxeles como blancos o negros según el criterio de un experto.

Después se compara la imagen binaria obtenida mediante SVM con la correspondiente *ground-truth*. En base a esto se calculan las siguientes cantidades:

- BV (blancos verdaderos): número de píxeles clasificados como blancos en la imagen SVM y también en la *ground-truth*.
- NV (negros verdaderos): número de píxeles clasificados como negros en la imagen SVM y también en la *ground-truth*.
- BF (blancos falsos): número de píxeles clasificados como blancos en la imagen SVM y negros en la *ground-truth*.
- NF (negros falsos): número de píxeles clasificados como negros en la imagen SVM y blancos en la *ground-truth*.

Con las siguientes cantidades se calcula el porcentaje de clasificación correcto (PCC) tal y como se describe la ecuación (6.2).

La tabla 6.2 muestra los valores PCC medios de las 40 imágenes analizadas para la identificación del verde y el tiempo medio de procesamiento según las ecuaciones (6.3) y (6.4) para los dos métodos utilizados, que en realidad se corresponden con el procesamiento de las imágenes en los modelos de color RGB y HSI.

Tabla 6.2 Valores PCC medios para la identificación del verdor en imágenes subexpuestas, normales (con y sin ajuste del TE) y sobreexpuestas; tiempo de procesamiento de cada método (μ s).

	Subexpuestas	Normal + ajuste			Sobreexpuestas
		no	RGB	HSI	
PCC (%)	66.4	87.3	91.1	92.2	45.1
Tiempo (μ s)			116	123	

Ambas estrategias de ajuste también han sido probadas con el tractor moviéndose en el campo y adquiriendo imágenes con una tasa de 10 imágenes por segundo, almacenando imágenes cada medio segundo para su posterior análisis. Del conjunto de 2547 imágenes disponibles, 1158 se adquirieron con el método RGB y 1389 con el método HSI. De ellas se seleccionan de forma aleatoria el 2% de las imágenes, es decir, 24 imágenes correspondientes al método RGB y 28 imágenes al HSI para las cuales se genera la imagen *ground-truth* de la ROI. El PCC medio alcanzado es 89.3 y 90.9 para cada estrategia RGB y HSI respectivamente, ambos valores se sitúan en el mismo orden de magnitud que los valores obtenidos previamente en la tabla 6.2.

De la tabla 6.2 se puede inferir que la peor situación se produce con imágenes sobreexpuestas, ya que en ese caso el sensor CCD es altamente impactado por la intensidad entrante, lo que imposibilita su recuperación. En cambio con las imágenes etiquetadas como subexpuestas sí es posible alcanzar ciertos resultados satisfactorios sin más que modificar las ganancias de la cámara o aplicando técnicas de ecualización sobre el histograma de la imagen. Ambas situaciones han de evitarse, pero en especial la sobreexposición.

6.3.3.2 Tendencia a adquirir imágenes subexpuestas, normales o sobreexpuestas

La segunda prueba se ha diseñado para estudiar el comportamiento de cada estrategia frente a su capacidad de adaptación a los cambios de iluminación. Para ello, el tractor se sitúa en la misma posición espacial dentro del campo, adquiriendo imágenes con ambas técnicas durante un día soleado con abundante nubosidad, de manera que existen cambios significativos y a veces bruscos en el nivel de iluminación. Las imágenes se almacenan junto con el tiempo de exposición aplicado en su captura, para el análisis posterior.

Como se puede ver en la figura 6.10, ambos métodos trabajan adecuadamente adquiriendo imágenes en su mayoría con un tiempo de exposición normal (entre 20 y 80 ms), aunque el método RGB tiende a capturar imágenes con un tiempo de exposición superior al requerido por el método HSI.

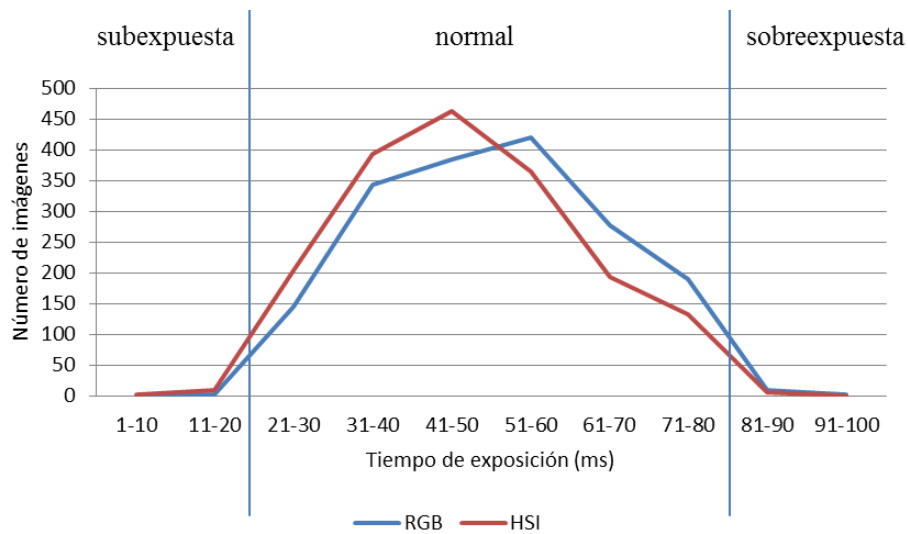


Figura 6.10 Variaciones del tiempo de exposición para los métodos RGB y HSI.

El hecho de que aparezcan algunas imágenes subexpuestas y sobreexpuestas se debe a que el algoritmo necesita de varias imágenes previas para calibrar el tiempo de exposición tras la puesta en marcha del sistema, siempre y cuando no se haya trabajado bajo las mismas condiciones de iluminación que la última vez que se fijó el tiempo de exposición. Como se comentó anteriormente, aunque se deben evitar las imágenes subexpuestas y sobreexpuestas, con las primeras es posible lograr resultados de segmentación aceptable. Así, se puede concluir que ambos métodos adquieren imágenes adecuadamente, si bien el método HSI supera al método RGB obteniendo valores inferiores en el tiempo de exposición.

6.3.4 Identificación de plantas y binarización basada en $YD_B D_R$

Las imágenes utilizadas para este estudio se corresponden con el conjunto descrito en la sección 6.3.2, considerando las correspondientes imágenes de *ground-truth*, generadas de la forma descrita en dicha sección. De las 200 imágenes disponibles se han seleccionado aleatoriamente 35 para este estudio. Las imágenes binarias obtenidas mediante la técnica $YD_B D_R$, descrita en el capítulo cuatro sección 4.3, se comparan con las correspondientes imágenes *ground-truth*. En base a todo lo anterior se calculan los siguientes valores:

- BV (blancos verdaderos): número de píxeles clasificados como blancos en la imagen binaria ($YD_B D_R$) y también en la *ground-truth*.
- NV (negros verdaderos): número de píxeles clasificados como negros en la imagen binaria ($YD_B D_R$) y también en la *ground-truth*.
- BF (blancos falsos): número de píxeles clasificados como blancos en la imagen binaria ($YD_B D_R$) y negros en la *ground-truth*.
- NF (negros falsos): número de píxeles clasificados como negros en la imagen binaria ($YD_B D_R$) y blancos en la *ground-truth*.

Con los valores anteriores se calcula el porcentaje de clasificación correcto (PCC) tal y como se describe la ecuación (6.2). El valor PCC medio de las 35 imágenes analizadas y el tiempo de procesamiento medio de cada una se muestran en la tabla 6.3.

Tabla 6.3 Valor PCC y tiempo de procesamiento medio para la segmentación basada en $YD_B D_R$.

	$YD_B D_R$
PCC (%)	88.6
Tiempo (ms)	93.2

A partir de los resultados mostrados en la tabla 6.3 se concluye que esta nueva estrategia de segmentación obtiene resultados buenos con un tiempo de procesamiento aceptable para trabajar en tiempo real. Por los resultados visuales, donde la figura 4.15 es una imagen representativa de este conjunto, se aprecia que esta estrategia provoca infra-segmentación en algunas imágenes, por lo tanto no requiere del uso de operaciones morfológicas con el fin de eliminar el ruido, lo cual puede ser una ventaja en función del resultado final deseado.

6.3.5 Segmentación e identificación de plantas enmascaradas y sin enmascarar mediante SVM

Las imágenes usadas para el estudio del comportamiento del método basado en SVM fueron adquiridas con la cámara digital HP Photosmart R817 en cuatro días diferentes durante abril/mayo de 2007. Todas las adquisiciones fueron espaciadas entre cinco y seis días. Un conjunto de 140 imágenes se tomaron en una fase de

pre-tratamiento, después el campo se regó artificialmente y recibió diferentes dosis de lluvia natural, las figuras 4.16 (a) y (b) son dos imágenes representativas de este conjunto. Un segundo conjunto de 70 imágenes se tomaron en una fase de post-tratamiento después de aplicar una dosis de herbicida, donde las malas hierbas habían empezado su proceso de secado. La figura 4.16(c) es una imagen representativa de este conjunto. Debido a la diferencia de días en las capturas, las imágenes se adquirieron también bajo distintas condiciones de iluminación. Estas circunstancias no afectan al rendimiento del proceso propuesto y por lo tanto no requiere ningún estudio adicional con respecto a la iluminación.

Estas imágenes se capturaron bajo una proyección de perspectiva que contiene sólo plantas y suelo, es decir, sin cielo. Las imágenes fueron capturadas en el espacio de color RGB con una representación de 24-bits por píxel y una resolución de 800x600 píxeles. Un total de 210 imágenes se procesaron con el Toolbox de Procesamiento de Imagen de Matlab R2009b en el portátil ASUS descrito previamente.

La estrategia propuesta se centra en la extracción de píxeles pertenecientes a plantas enmascaradas y sin enmascarar o píxeles afectados por el tratamiento. El procedimiento se basa en la estrategia de aprendizaje descrita en el capítulo cuatro sección 4.4.3. Del anterior conjunto disponible de 210 imágenes se seleccionaron aleatoriamente 70 para identificar los vectores soporte, que constituyen el conjunto de entrenamiento.

Con respecto a la elección del núcleo H presente en las ecuaciones (4.18) y (4.19), se seleccionaron aleatoriamente cinco imágenes del conjunto de 70 utilizadas en el aprendizaje. Manualmente se seleccionaron 1000 píxeles, 500 marcados como $y_k = +1$ (clase C_1) y 500 marcados como $y_k = -1$ (clase C_2). Aleatoriamente se selecciona el 80% de los píxeles de cada conjunto de 500, que se usaron como patrones de entrenamiento, y a partir de ellos se obtiene la función de decisión en (4.18) y el parámetro b en (4.23). El 20% restante de los patrones seleccionados se usaron para la validación. Así, dado un patrón x , éste se asigna a la clase C_1 si $f(x) > 0$ o a la clase C_2 si $f(x) \leq 0$. Se han analizado los siguientes tres núcleos:

- a) *Función Gaussiana de base radial*: $H(x, y) = \exp\{-\|x - y\|/\sigma^2\}$, donde se varía σ^2 de 1 a 10 en pasos de 0.5 y para cada σ^2 se calcula el porcentaje de éxito obtenido con los patrones de validación de datos. Este proceso se repite 10 veces. Finalmente se obtiene $\sigma^2 = 3.5$ con una media de porcentaje de éxito del 84.2%.
- b) *Polinomio*: $H(x, y) = \langle x, y \rangle^d$, con $\langle x, y \rangle$ definido como el producto escalar. Siguiendo el mismo procedimiento que el anterior, pero variando el parámetro d de 1 a 5 en pasos de 0.5, se obtiene $d = 2.5$ con un porcentaje de éxito del 79.1%.
- c) *Sigmoide*: $H(x, y) = \tanh(\rho \langle x, y \rangle + \gamma)$, como las anteriores, se varía ρ de 1 a 10 en pasos de 0.5 y γ de 1 a 5 en pasos 0.5 obteniendo $\rho = 2.5$ y $\gamma = 3.0$ con un 74% de porcentaje de éxito.

Finalmente se elige la función Gaussiana de base radial debido a su rendimiento con $\sigma^2 = 3.5$.

Tal y como se describe en el capítulo cuatro sección 4.4.3, dada una imagen, la imagen binaria se obtiene a través de la estrategia combinada COM, ecuación (4.17), de los tres índices vegetativos seguida de una umbralización mediante Otsu. Las figuras 4.19 (a) y (b) muestran respectivamente la imagen COM y el resultado después de aplicar Otsu a la imagen original de la figura 4.16(a), que resulta ser una imagen representativa del conjunto de imágenes utilizadas.

La figura 4.18 muestra la distribución de píxeles, en el espacio de color tridimensional RGB pertenecientes al conjunto de entrenamiento. Los píxeles marcados en verde son los patrones pertenecientes a la clase C_1 , es decir, los píxeles blancos en las correspondientes imágenes binarias. Por el contrario, los píxeles marcados en rojo representan los patrones pertenecientes a la clase C_2 , píxeles negros en las imágenes binarias. El número de píxeles usado para el entrenamiento fue 800 (alto de la imagen) \times 600 (ancho de la imagen) \times 70 (número de imágenes). En promedio, el porcentaje de patrones pertenecientes a las clases C_1 y C_2 fue respectivamente 35% y 75%. La figura 4.18 también muestra ambos tipos de vectores soporte rodeados por dos elipses,

los puntos negros representan los vectores soporte asociados a la clase C_1 , es decir, pertenecientes a plantas sin enmascarar; los puntos azules se asocian a las plantas enmascaradas.

Como se mencionó en el capítulo cuatro sección 4.4.3, el marco teórico de las SVM establece estrictamente que los vectores soporte están asociados con un α_k distinto de cero, en los experimentos se ha descartado también valores de α_k donde los valores absolutos están por debajo de un cierto umbral que se ha fijado a 10^{-3} . En promedio, sobre el total de imágenes usadas para el entrenamiento, el número de vectores soporte por imagen fue de 106, de los que un 60% estaban asociados a la clase C_1 y un 40% a la clase C_2 .

La tabla 6.4 muestra los valores para los tres parámetros (w , τ , b) calculados con las ecuaciones (4.21), (4.22) y (4.23) sobre un conjunto de 7420 vectores soporte extraídos del conjunto de 70 imágenes.

Tabla 6.4 Número de patrones procesados, vectores soporte obtenidos y parámetros obtenidos basados en el marco de los SVM.

Número de patrones (píxeles) procesados	Número de vectores soporte	w	τ	b
$800 \times 600 \times 70$	7420	$(3.3, 3.4, 2.5) \times 10^4$	1.86×10^{-5}	-1.2×10^2

A partir de los resultados mostrados en la tabla 6.4 cabe resaltar lo siguiente: 1) el número de vectores soporte obtenidos por las SVM representa sólo el 0.022% del total de patrones procesados y 2) el margen τ tiene un valor muy pequeño, y como éste representa la distancia mínima de los hiperplanos de separación a los puntos de datos, cabe concluir que los puntos de datos se sitúan muy próximos al hiperplano y por consiguiente las clases C_1 y C_2 están en cierto modo solapadas. Sin embargo, la estrategia propuesta es capaz de distinguir entre los dos tipos de plantas, tal y como se esperaba.

La tabla 6.5 muestra los valores resultantes de los parámetros de los vectores soporte para las clases C_1 y C_2 durante la fase de aprendizaje. Estos son valores promedios para las componentes espectrales verdes (\bar{x}_u, \bar{x}_m) , desviaciones estándar

(σ_u, σ_m) , radios (r_{uG}, r_{mG}) y tolerancias de verde (t_{uG}, t_{mG}) , todos ellos incluidos en las ecuaciones (4.24), (4.25) y (4.26).

Tabla 6.5 Valores de los parámetros para la fase de decisión durante la identificación de plantas.

C_1			
$\bar{x}_u = (\bar{R}_u, \bar{G}_u, \bar{B}_u)$	σ_u	r_{uG}	t_{uG}
(160.5, 165.5, 108.7)	(18.1, 16.3, 20.2)	0.3809	0.0984
C_2			
$\bar{x}_m = (\bar{R}_m, \bar{G}_m, \bar{B}_m)$	σ_m	r_{mG}	t_{mG}
(102.2, 100.8, 84.4)	(9.3, 10.3, 9.5)	0.3507	0.1022

A partir de los resultados mostrados en la tabla 6.5 se observa cómo la componente espectral verde en C_1 es mayor que la roja. Esto establece el límite inferior, que resulta ser suficiente para identificar plantas sin enmascarar caracterizadas por este comportamiento. Esto no ocurre en C_2 donde la componente espectral roja es la dominante, aunque sea en una pequeña cantidad; esto se debe a la reflectancia producida por los elementos del barro proveniente del suelo, los cuales han contaminado parte de las plantas originales, o bien a que las plantas han iniciado su proceso de secado debido al tratamiento con herbicida. También se puede observar la pequeña diferencia relativa entre los radios r_{uG} y r_{mG} con respecto a los márgenes de tolerancia expresados mediante t_{uG} y t_{mG} . La conclusión final es que bajo el paradigma de las SVM es posible distinguir entre plantas sin enmascarar y enmascaradas, en este último caso plantas afectadas por el tratamiento o por haber sido impregnadas con partículas arcillosas procedentes del suelo.

La figura 4.20(a) muestra la imagen final segmentada a partir de la imagen de la figura 4.16(b), representativa del conjunto de imágenes utilizadas, donde las marcas verdes y rojas insertadas sobre la imagen original corresponden a plantas sin enmascarar y enmascaradas respectivamente. Se puede observar cómo la mayoría de marcas verdes están asociadas al cultivo y la mayoría de las marcas rojas pertenecen a malas hierbas o a cultivo muy próximo al suelo, y por tanto enmascarando los verdes. En el primer caso, las plantas de maíz no han sido impregnadas, entre otras razones porque su altura con respecto al suelo lo ha evitado. Esta circunstancia también aparece en algunas malas hierbas, donde como antes, su altura evita esa impregnación.

En relación a las imágenes disponibles para el análisis post-tratamiento, representadas por la imagen de la figura 4.20(b), se puede observar a partir de esta imagen cómo las partes más cercanas al suelo aparecen como enmascaradas, y también cómo una importante cantidad de malas hierbas, en la zona central de las líneas de cultivo, aparecen altamente afectadas por el tratamiento, lo que significa que han empezado su proceso de secado.

Adicionalmente, y con el fin de evaluar la validez de la estrategia propuesta, se seleccionaron aleatoriamente 40 imágenes originales, que fueron analizadas por un experto para la identificación de malas hierbas y cultivo. La observación visual se llevó a cabo para cada imagen usando como referencia la imagen segmentada a través del enfoque propuesto en el capítulo cuatro, sección 4.4. El experto concentró su mayor esfuerzo en la identificación de las plantas problemáticas, es decir, sobre las denominadas plantas enmascaradas. Las asignaciones incorrectas se marcaron manualmente, se corrigieron o se eliminaron, generando una nueva imagen segmentada, considerada como imagen de referencia (o *ground-truth*). En promedio, sobre el conjunto de imágenes utilizadas se obtuvo un porcentaje de éxito del 93.1%.

6.4 Análisis del sistema experto

En esta sección se estudian las técnicas propuestas a lo largo del capítulo cinco para la detección de líneas de cultivo y malas hierbas, guiado del tractor autónomo y solapamiento de las zonas de tratamiento específico.

6.4.1 Estimación de líneas de cultivo: detección de la orientación

Las imágenes utilizadas en este estudio pertenecen a campos de maíz. Han sido tomadas con la cámara a color SVS-VISTEK SVS4050CFLGEA durante los meses de marzo/mayo de 2012, 2013 y 2014. Las imágenes digitales fueron tomadas bajo proyección de perspectiva y almacenadas como imágenes con resoluciones de 2336x1752 píxeles a color de 24 bits, guardadas en formato BMP con el modelo de color RGB. En este caso, los parámetros extrínsecos e intrínsecos del sistema de visión vienen dados por la configuración necesaria para desempeñar su función dentro del proyecto RHEA. Dado que la cámara ha sido instalada en el tractor tal y como se puede ver en la figura 1.2 y siguiendo los requisitos establecidos en el proyecto para la

detección de malas hierbas en cuatro líneas de cultivo, los parámetros son: número de líneas de cultivo = 4, ángulo de cabeceo (*pitch*) = 22° , ángulo de alabeo (*roll*) = 0° y ángulo de guiñada (*yaw*) = 0° con una altura de la cámara con respecto al suelo de 2.23 metros y una distancia focal de 10.4 milímetros. La ROI es la definida en el capítulo cuatro, sección 4.2.2.2. En este estudio, se seleccionaron aleatoriamente 300 imágenes del total de más de 25000 imágenes disponibles con distintas orientaciones del tractor en el campo con respecto a las líneas de cultivo reales. Las imágenes obtenidas abarcan un amplio abanico de orientaciones como puede observarse en los ejemplos mostrados en las figuras 5.3(a) y 5.5(a) respectivamente.

El estudio se ha llevado a cabo para diferentes alturas de las plantas de maíz, agrupadas en tres clases: *baja*, para plantas entre 1–4 centímetros, *media*, entre 4–8 centímetros y *alta* para más de 8 centímetros. Además, se ha limitado la búsqueda del ángulo de guiñada, según se describe en el capítulo cinco sección 5.4.2, para distintos rangos, donde 0° implica no aplicar el algoritmo de detección de la orientación, asumiendo que el tractor está bien orientado. Suponiendo que las líneas de cultivo presentan un ancho para el cual las plantas representan el cultivo, el ancho aplicado a cada línea estimada se fija en el 18% del ancho entre dos líneas de cultivo adyacentes. Dado que las líneas han sido sembradas con una distancia de 75 cm entre ellas, este porcentaje equivale a 13.5 cm a cada lado de la recta estimada.

Además de los ángulos fijos descritos anteriormente se ha utilizado la IMU MicroStrain 3DM-GX3-35 descrita en la sección 6.2.3 para obtener los ángulos de cabeceo y alabeo con los que se ha capturado cada imagen. La finalidad de utilizar la IMU es estudiar cómo afecta la variabilidad del terreno a la precisión cuando se utilizan ángulos fijos y variables.

La tabla 6.6 muestra los porcentajes medios de éxito junto con los tiempos de procesamiento para los tres estados de crecimiento de las plantas y tres valores de ángulos de guiñada. Los porcentajes de éxito se determinan analizando para cada imagen el grado de coincidencia entre la línea estimada y la que teóricamente debería ser, atendiendo al criterio de un experto.

Tabla 6.6 Porcentaje de acierto medio y tiempos de procesamiento (ms) de imágenes capturadas con la cámara SVS-VISTEK SVS4050CFLGEA para la detección de líneas de cultivo con diferentes alturas de plantas y rangos de búsqueda del ángulo de guiñada.

	Altura Giro máx.	baja			media			alta		
		0°	5°	10°	0°	5°	10°	0°	5°	10°
Porcentaje de éxito	Áng. fijos	73	88	93	67	86	84	62	81	77
	IMU	77	91	97	70	90	88	65	85	81
Tiempo (ms)	Áng. fijos	13	21	27	13	22	27	14	22	28
	IMU	13	22	29	14	22	29	14	23	30

En los resultados mostrados en la tabla 6.6, se puede observar que cuanto mayor es la altura de las plantas peor es el porcentaje de éxito. Esto se debe a que las líneas de cultivo poseen una mayor cantidad de píxeles verdes y por lo tanto es más complicado que el algoritmo detecte la nube de puntos que mejor aproxima a las líneas de cultivo. Además, debido a las diferencias en el crecimiento, las líneas presentan mayores irregularidades.

En cuanto al giro máximo aplicado según las diferentes alturas se puede observar cómo sin realizar giro (0°) los resultados no son satisfactorios ya que el tractor ha de estar perfectamente orientado para que el resultado sea óptimo. En cambio, cuando se aumenta este rango se dan dos situaciones. Por un lado, cuando la altura es baja, a mayor ángulo mejores resultados se obtienen debido a que las líneas están más acotadas y por lo tanto la búsqueda de verdes es más precisa. Por otro lado, se observa cómo cuando aumenta la altura de las plantas, el hecho de aumentar el ángulo implica que los resultados sean ligeramente peores. Esto se debe en gran medida a que la cantidad de píxeles verdes aumenta y por lo tanto son mayores las posiciones donde aparecen líneas candidatas, lo que genera distintas soluciones sin poder concretar cuál de ellas es la que mejor aproxima a todas las líneas de cultivo. En el caso de utilizar los ángulos proporcionados por la IMU, la ligera mejoría es consecuencia exclusiva de los ángulos de cabeceo y alabeo, ya que cuando su valor es fijo no se tiene en cuenta la inclinación sufrida por el sistema de visión, y aunque por lo general el campo es un terreno llano, en algunas ocasiones el tractor sufre distintas inclinaciones debido a las rodadas de la maquinaria y al propio terreno que sufre altibajos.

En cuanto al tiempo de procesado, cabe destacar que por regla general, cuanto mayor es el ángulo de guiñada mayor es el tiempo de cómputo, ya que la zona de exploración aumenta al existir más posiciones posibles. Las variaciones sufridas, pese a ser realmente despreciables, se deben a pequeños retardos a la hora de establecer la comunicación con la IMU y al número de operaciones internas que ha de realizar, ya que a mayor cantidad de verdes más veces hay que incrementar la variable asignada a la detección del número de verdes por línea.

6.4.2 Corrección de líneas mediante el estimador Theil-Sen

Las imágenes utilizadas en este estudio pertenecen a campos de maíz tomadas con la cámara a color BASLER scA1400-17FC, bajo proyección de perspectiva, durante los meses de abril/mayo de 2011 y almacenadas como imágenes con resoluciones de 1392x1038 píxeles a color de 24 bits, guardadas en formato TIFF con el modelo de color RGB. La adquisición se llevó a cabo durante tres semanas distintas en un campo de 1.7 hectáreas situado en la Finca Experimental del CSIC en La Poveda (Arganda del Rey, Madrid), por lo que se obtuvieron imágenes con distintas condiciones de iluminación y diferentes estados de crecimiento tanto de las plantas de maíz como de las malas hierbas.

Las imágenes fueron procesadas con LabVIEW Real-Time de National Instruments (2014), edición 2011, bajo la cRIO-9082 con Sistema Operativo en Tiempo Real (*Real-Time Operating System*). El algoritmo propuesto está desarrollado en C mediante MS Visual Studio y compilado como una DLL, que se ha embebido en un módulo adicional en LabVIEW. Se ha procesado un conjunto de 240 imágenes. Además, este equipamiento está diseñado para cumplir con las especificaciones en tiempo real que aparecen en el punto g) del capítulo cinco sección 5.2.

Los parámetros extrínsecos e intrínsecos de la cámara son: número de líneas de cultivo = 3, ángulo de cabeceo (*pitch*) = 20°, ángulo de alabeo (*roll*) = 0° y ángulo de guiñada (*yaw*) = 0° con una altura de la cámara con respecto al suelo de 1.5 m, una distancia focal de 8.5 mm y la zona de estudio situada a 2 m con respecto a la base de la cámara, extendiéndose hasta los 6 m.

La figura 6.11, muestra un resultado ilustrativo sobre la mejora del estimador Theil-Sen propuesto comparado con el método de ajuste basado en regresión lineal simple.

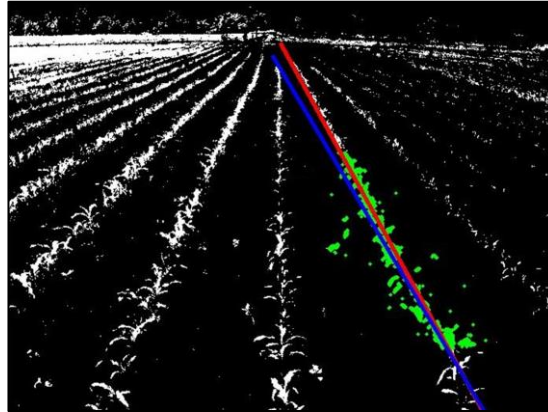


Figura 6.11 Ejemplo de corrección de líneas de cultivo mediante el estimador Theil-Sen (línea roja) y la regresión lineal simple (línea azul) considerando la dispersión de plantas (píxeles verdes).

Los píxeles verdes de la línea de cultivo derecha son píxeles marcados que se han obtenido considerando la línea de cultivo estimada mediante la aplicación de las transformaciones geométricas de acuerdo a los parámetros intrínsecos y extrínsecos del sistema de visión, y aplicando un margen de tolerancia fijado al 50% (37.5 cm a cada lado) tal y como se ha explicado en el capítulo cinco sección 5.4. Como se deduce fácilmente, los píxeles más alejados del centro corresponden a malas hierbas, es decir, aparecen dispersos en la zona que separa las líneas de cultivo. La línea roja en la imagen representa la línea de cultivo corregida aplicando el estimador Theil-Sen y la línea azul la correspondiente a aplicar una regresión lineal simple. Como se puede observar, el mejor ajuste lo realiza el método de Theil-Sen. Esto se debe a que es lo suficientemente robusto como para hacer frente a la dispersión de píxeles. Por el contrario, el método basado en la regresión es sensible a este tipo de dispersiones ya que se basa en el cálculo de distancias mínimas y los píxeles aislados influyen atrayendo el ajuste.

El comportamiento del estimador Theil-Sen con respecto a la regresión lineal simple se estudia a través de un análisis cualitativo basado en el criterio de un experto ya que no hay imágenes de referencia (*ground-truth*). Mediante una inspección visual de las líneas de cultivo reales en las imágenes, el experto determina el mejor ajuste de las líneas de cultivo corregidas.

Debido a que este enfoque está orientado a trabajar en aplicaciones de tiempo real, una vez se ha conseguido un funcionamiento satisfactorio en la detección de líneas de cultivo, se analiza el comportamiento de ambos estimadores cuando se aplican a imágenes con diferentes resoluciones, dando lugar a un análisis cuantitativo. A continuación se analizan los resultados tanto cualitativamente como cuantitativamente.

La figura 6.12 (a) y (b) muestra a modo de ejemplo una de las 240 imágenes analizadas. El área de interés está delimitada por unas líneas de referencia con un ancho de 2.25 m (abarcando 3 líneas de cultivo espaciadas cada una 0.75 m) y una longitud de 4 m. En este tipo de imágenes esta área representa la ROI objeto del tratamiento y por tanto del procesamiento.

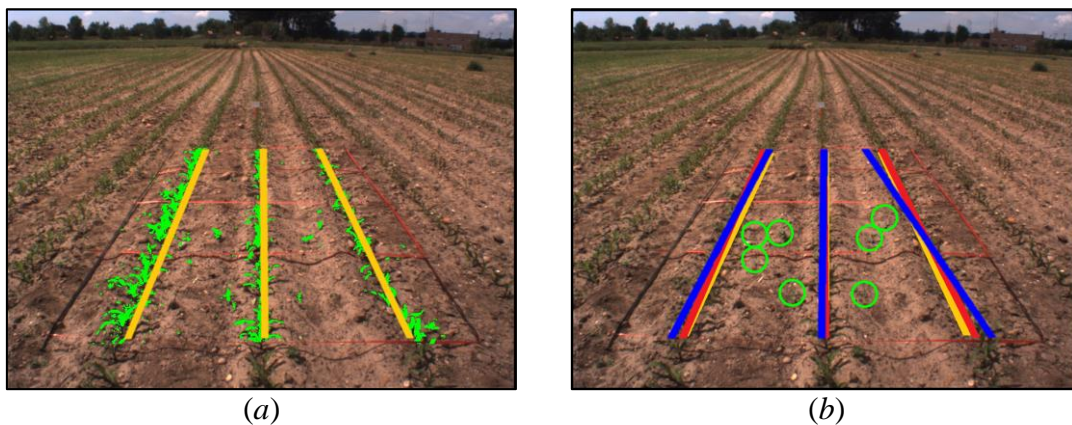


Figura 6.12 (a) Líneas de cultivo idealmente estimadas calculadas mediante los parámetros intrínsecos y extrínsecos (líneas amarillas) y un margen de tolerancia del 50%. (b) Corrección de las líneas usando el estimador Theil-Sen (líneas rojas) y la regresión lineal simple (líneas azules). Las circunferencias verdes marcan las zonas con píxeles de malas hierbas.

6.4.2.1 Análisis cualitativo

La figura 6.12(a) muestra en verde aquellos píxeles que representan plantas verdes (cultivo y malas hierbas). Las líneas de cultivo estimadas con respecto al sistema de geometría basado en los parámetros extrínsecos e intrínsecos se dibujan como líneas amarillas. Una inspección simple de la imagen bajo el criterio de un experto permite deducir que las líneas de cultivo estimadas no coinciden con las reales con la precisión deseada, particularmente las exteriores. Esta situación es especialmente relevante en la línea de cultivo izquierda. Esto se debe a que dicha línea de cultivo no se sembró con la separación de los 0.75 m requeridos respecto del espaciado interlineal teórico en campos de maíz. La desviación con respecto a la línea de cultivo real de las líneas de cultivo estimadas tanto central como derecha es menor que en el caso de la de la

izquierda. Es obvio que bajo esta situación las líneas de cultivo necesitan una corrección. Las tres líneas de cultivo reales contienen algunos huecos producidos por errores en la siembra.

La figura 6.12(b) muestra las correcciones de las líneas amarillas aplicando tanto el estimador Theil-Sen (líneas rojas) como la regresión lineal simple (líneas azules). Basándose en la experiencia humana, es fácil deducir cómo las tres líneas rojas aparecen ajustadas de forma satisfactoria por el método de Theil-Sen, de hecho siguen la parte central de la línea de cultivo. Quizás se puede apreciar una ligera desviación en la línea central, aunque menor que la desviación que produce la regresión lineal. La acusada desviación producida por este último estimador en la línea de cultivo derecha se debe a la presencia de rodales aislados (marcados con circunferencias verdes), correspondientes a malas hierbas, que han intervenido en el proceso de ajuste. Algo similar ocurre en la línea de cultivo izquierda, aunque la desviación es menos pronunciada.

Un comportamiento similar se observó en el conjunto de las 240 imágenes analizadas, lo cual nos permite concluir que el estimador Theil-Sen funciona mejor que la regresión lineal simple. El margen de tolerancia usado fue de nuevo del 50% de la distancia entre líneas de cultivo a cada lado, es decir, una distancia de 37.5 centímetros a cada lado de la recta estimada. En definitiva el método de Theil-Sen funciona correctamente bajo éste tipo de situaciones, donde existen malas hierbas cercanas al cultivo y con errores en la siembra, las cuales son abundantes en los campos de maíz.

Con el fin de detectar las líneas de cultivo reales sobre la imagen, también se ha aplicado la transformada de Hough (Slaughter y col., 2008). A pesar de que se utilizan restricciones geométricas considerando los parámetros extrínsecos e intrínsecos, el método produce abundantes máximos en las celdas acumuladoras, haciendo difícil la selección del máximo que identifica la línea de cultivo. Aparecen muchas líneas con diferentes pendientes para una única línea de cultivo, lo que requiere una cuidadosa selección del umbral máximo (Jones y col., 2009a, b; Rovira-Más y col. 2005). Además, los tiempos requeridos para el procesamiento mediante la transformada de Hough, resultan ser más elevados que en el caso de los estimadores estudiados, razón por la cual se descartó como método aplicable para tiempo real.

6.4.2.2 Análisis cuantitativo

Desde el punto de vista de un análisis cuantitativo se estudia el porcentaje de acierto medio basándose en el criterio de un experto tanto para la regresión lineal simple como para Theil-Sen, así como sin ajuste. Cada imagen es visualmente analizada por el experto para identificar la precisión entre la línea ajustada y la línea de cultivo real que el experto considera como satisfactoria. Se calcula el coste computacional midiendo el tiempo de procesamiento para cada una de las tres estrategias de ajuste. Se denomina sin ajuste (SA) al proceso que envuelve la extracción de verde mediante el método GA descrito en el capítulo cinco sección 5.3 apartado *b*), en este caso la binarización se realiza mediante el método Otsu, tras el cual se aplica la geometría del sistema para trazar las líneas esperadas sobre la imagen basándose en los parámetros intrínsecos y extrínsecos y suponiendo una buena orientación del sistema con respecto al campo. La tabla 6.7 muestra los porcentajes de éxito para cada uno de los métodos (SA, Theil-Sen y Regresión) para distintos tamaños de imagen y márgenes de búsqueda para el ajuste, como se especifica más adelante.

En relación a los tiempos de cómputo, en la tabla 6.7 se muestran los obtenidos para SA, mientras que para los métodos de Regresión y Theil-Sen, los tiempos que aparecen son exclusivamente los que se obtienen para llevar a cabo el ajuste específico en cada caso. El coste total en estos dos métodos es el resultado de sumar el coste sin ajuste (SA) más los que aparecen en la tabla.

Debido a que el método de Theil-Sen es, en términos computacionales, ligeramente elevado pero efectivo, con el fin de analizar más a fondo su comportamiento se reduce la ROI de la imagen original submuestreando la correspondiente ROI de la imagen original hasta obtener reducciones equivalentes al 30% y 60% de la imagen original con el fin de comprobar su rendimiento con y sin reducciones. La tabla 6.7 muestra los porcentajes medios y los tiempos computacionales expresados en milisegundos para el conjunto de imágenes originales disponibles (es decir, con un 0% de reducción) y también para dichas imágenes cuando la ROI es reducida al 30% y 60% respectivamente, tanto el tamaño horizontal como vertical. Para cada una de esas reducciones se muestra el porcentaje medio de aciertos junto con los tiempos computacionales para tres valores (5%, 18% y 50%) de lo que se denomina márgenes de tolerancia. Dichos márgenes representan, en porcentaje, el ancho a cada

lado de las líneas de cultivo estimadas sin ajuste en relación a la separación de dos líneas adyacentes, siendo equivalentes a 3.75 cm, 13.5 cm y 37.5 cm respectivamente sobre el terreno real, considerando que la separación real entre líneas de cultivo es de 75 cm. Dentro de dichos márgenes aparecen tanto píxeles relativos al cultivo como a las malas hierbas alrededor de la línea de cultivo estimada directamente por aplicación de las restricciones geométricas, con los parámetros extrínsecos e intrínsecos, sin ajuste. Bajo las consideraciones anteriores, todos los valores mostrados en la tabla 6.7 para SA no varían, ya que en este caso no es necesario realizar búsqueda alguna dentro de ningún margen.

Tabla 6.7 Porcentaje de acierto medio y tiempos de procesamiento (ms) de imágenes capturadas con la cámara BASLER scA1400-17FC para SA, regresión lineal y Theil-Sen con diferentes reducciones de tamaño y márgenes de tolerancia.

	Reducción Márgenes	0%			30%			60%		
		5%	18%	50%	5%	18%	50%	5%	18%	50%
Porcentaje de éxito	SA	75	75	75	75	75	75	75	75	75
	Regresión	85.1	89.2	86.3	83.9	88.5	85.2	80.2	84.2	76.1
	Theil-Sen	93.3	95.6	94.2	90.6	95.2	93.7	88.3	93.8	92.1
Tiempo (ms)	SA	146	146	146	87	87	87	65	65	65
	Regresión	163	174	186	97	95	106	74	83	91
	Theil-Sen	339	380	425	253	264	285	241	249	257

A partir de los resultados mostrados en la tabla 6.7 se observa que el mejor rendimiento, en términos de precisión, se consigue mediante el método de Theil-Sen comparándolo con el método de regresión lineal simple y SA. Por otro lado, el mejor rendimiento absoluto de Theil-Sen, también en términos de precisión, se obtiene utilizando la ROI original y un margen de tolerancia del 18%. Esto se debe a que esas imágenes contienen el máximo de información de la zona de interés y el margen cubre el ancho necesario para recopilar toda la información proveniente de las líneas de cultivo sin que los píxeles de malas hierbas influyan excesivamente sobre el ajuste. Cuando la reducción se incrementa la precisión disminuye. Esta disminución se explica por el hecho de que a mayor reducción y tolerancia, proporcionalmente más píxeles pertenecientes a malas hierbas intervienen en la estimación de las líneas. En lo que se refiere a la tolerancia, un margen elevado de ésta implica que más píxeles pertenecientes a malas hierbas intervienen en la estimación de las líneas; por el contrario, un número reducido impide una correcta identificación cuando la geometría

no es precisa, ya que pierde la información necesaria perteneciente a las líneas de cultivo. Por este motivo, los mejores resultados en términos de porcentaje de éxito se obtienen con un valor de tolerancia intermedio, como en este caso de 18%.

Se han probado diferentes combinaciones de márgenes y reducciones sin una mejoría aparente respecto a las mostradas en la tabla 6.7.

En cuanto a los tiempos, es obvio que cuanto mayor sea el tamaño de la imagen y los márgenes de tolerancia, mayores son los tiempos de procesado, ya que es necesario procesar un mayor número de píxeles. El principal inconveniente para Theil-Sen es su mayor coste computacional en comparación con los otros dos métodos. No obstante, asumiendo que el sistema de visión procesa un área de 4 metros de longitud, el tractor necesitará navegar a velocidades inferiores a $4/T$ (m/s) para que dé tiempo a que se procese la imagen y se actúe, donde T representa el tiempo de procesado. Así, dado que en el peor de los casos el tiempo resulta ser $T = 571$ ms (146 ms de SA + 425 ms de Theil-Sen), la velocidad máxima a la que debería ir el tractor sería de 25.21 km/h, velocidad claramente excesiva para tareas de tratamiento. Esto significa que es innecesario realizar reducciones de tamaño sobre la ROI ya que en el peor de los casos el sistema de visión garantiza tiempos de cómputo acordes con las velocidades de trabajo normales, que oscilan entre 3 y 6 km/h para un tratamiento térmico, y entre 6 y 8 km/h para la aplicación de herbicida con una barra de pulverización.

Para clarificar los resultados de tiempo mostrados en la tabla 6.7, en la figura 6.13 se muestran los tiempos medios de procesado, en ms, sobre el conjunto de imágenes disponibles para el proceso completo, incluyendo la extracción de verdes mediante el método GA, la binarización mediante el método de Otsu y el ajuste de líneas basado en el estimador de Theil-Sen. Como se indicó anteriormente, estos tiempos se obtienen para tres valores de reducción (0%, 30% y 60%) y con tres márgenes de tolerancia (5%, 18% y 50%). Gráficamente se observa cómo el tiempo de procesamiento decrece con la reducción y con el margen de tolerancia.

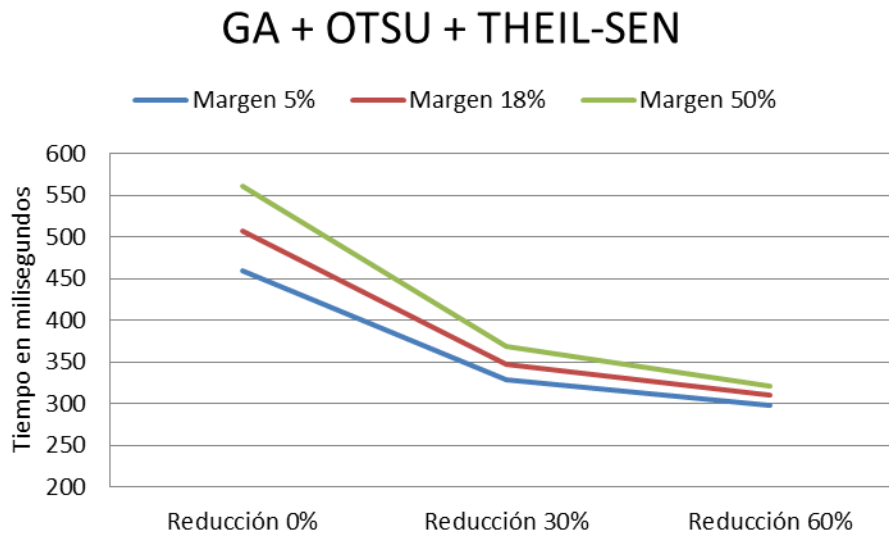


Figura 6.13 Promedio de los tiempos de procesamiento para el proceso completo: extracción de verdes (GA), binarización (Otsu) y ajuste de líneas (Theil-Sen) para los tres tamaños de reducción y los tres márgenes de tolerancia.

Con el fin de validar esta estrategia con imágenes obtenidas con otras cámaras, se ha aplicado esta técnica al conjunto de imágenes descrito en la sección 6.4.1 tras haber detectado la orientación correcta. Al tratarse de cámaras diferentes a las utilizadas aquí, proporcionan imágenes con unas características diferentes. Además, la configuración del sistema de visión es completamente distinta, tanto en lo que se refiere a la altura como en relación a los ángulos de inclinación, así como en las especificaciones de la ROI y el número de líneas de cultivo. Así, queda demostrada la generalidad de la aproximación propuesta para trabajar sobre diferentes sistemas de visión, ya que sólo es necesario configurar convenientemente un conjunto de parámetros para obtener buenos resultados.

La figura 5.8 muestra dos imágenes representativas de este conjunto de imágenes donde se puede ver cómo el estimador Theil-Sen resulta suficientemente robusto frente a distintas configuraciones del sistema siendo capaz de hacer frente incluso a fuertes desviaciones respecto de la estimación inicial.

La tabla 6.8 muestra los porcentajes de éxito medios y los tiempos computacionales expresados en milisegundos para el conjunto de imágenes originales disponibles (es decir, con un 0% de reducción) y también los distintos márgenes de tolerancia (5%, 18% y 50%) que, al igual que en las imágenes anteriores, son equivalentes a 3.75 cm, 13.5 cm y 37.5 cm respectivamente.

Tabla 6.8 Porcentaje de acierto medio y tiempos de procesamiento (ms) de imágenes capturadas con la cámara SVS-VISTEK SVS4050CFLGEA para SA, regresión lineal y Theil-Sen con diferentes márgenes de tolerancia.

	Márgenes	5%	18%	50%
Porcentaje de éxito	SA	72	72	72
	Regresión	82.5	87.3	84.5
	Theil-Sen	91.1	93.8	92.2
Tiempo (ms)	SA	162	162	162
	Regresión	188	197	204
	Theil-Sen	391	434	479

Con esta cámara en el peor de los casos el tiempo sería $T = 641$ milisegundos (162 ms de SA + 479 ms de Theil-Sen), por lo que la velocidad máxima a la que podría navegar el tractor sería de 11.23 km/h, siendo una velocidad menor que en el caso anterior pero igualmente alta para tareas de tratamiento.

Por lo tanto, queda puesto de manifiesto la viabilidad de esta técnica para aplicaciones agrícolas que requieran la identificación precisa y en tiempo real de líneas de cultivo.

6.4.3 Guiado del tractor

Para el estudio de las técnicas de guiado se han realizado diferentes trayectorias de forma manual georreferenciando los puntos por los que debe pasar el tractor. Con esto se consigue disponer de las posiciones concretas (trayectoria) por las que ha de pasar el tractor durante el tratamiento específico por zonas.

Una vez se dispone de estos valores se ejecutan diferentes trayectorias teniendo en cuenta la recta generada por el planificador en función de los puntos de entrada y salida del campo proporcionados por el GPS. Esta recta ideal se ve modificada por aplicación de las técnicas de corrección descritas en el capítulo cinco sección 5.6, almacenándose las coordenadas GPS por las que pasa realmente el tractor en cada trayectoria. El objetivo es calcular la desviación estándar de cada una de las coordenadas según la ecuación (6.5).

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\left(\sqrt{(x_i - X_i)^2 + (y_i - Y_i)^2} \right) - \bar{d} \right)^2} \quad (6.5)$$

donde N es el número total de puntos georreferenciados, \bar{d} es la media del total de distancias y tanto (x, y) como (X, Y) son las coordenadas georreferenciadas en la posición i cuando el tractor va de forma automática y manual respectivamente.

Tras realizar 20 trayectorias con diferentes configuraciones tanto del sistema de visión como de los puntos de entrada y salida del campo, con una altura de maíz baja (entre 1–4 centímetros) se ha obtenido una desviación estándar σ de 7.7 centímetros, mientras que con una altura superior la desviación aumenta a 10.2 centímetros debido a la imprecisión a la hora de detectar las líneas de cultivo. Debido a estos errores, se ha decidido realizar la corrección siempre que el error se sitúe entre 2–6 centímetros, obviando las correcciones de menor cantidad ya que se supone que la trayectoria que se está siguiendo es suficientemente buena. Se suprimen las correcciones por encima de tales valores porque seguramente el cálculo provenga de un error en la identificación de las líneas de cultivo.

También se han realizado pruebas donde se han desplazado los puntos de salida del campo, de manera que la trayectoria de referencia generada (puntos GPS) no es una línea paralela al cultivo, observándose que la técnica de guiado propuesta responde adecuadamente a desplazamientos de hasta 40 centímetros.

6.4.4 Solapamiento de imágenes

Para el análisis del solapamiento se han georreferenciado las ROIs en el campo y marcado con el fin de comprobar visualmente que las proyecciones corresponden con el inicio y fin de cada ROI. Para cada ROI se ha posicionado el tractor con la configuración descrita en la sección 6.4.1 a la distancia exacta de 3.3 metros y se han obtenido las coordenadas GPS de cada una de esas posiciones.

Con todos los valores se realizan diferentes trayectorias teniendo en cuenta la recta generada por el planificador en función de las coordenadas GPS de los puntos de entrada y salida del campo. Cada vez que se toma una nueva imagen, las coordenadas GPS del tractor se comparan con las de la última captura obteniendo así la distancia a la que se ha adquirido esa imagen según la ecuación (5.10). La distancia a la que debe coincidir idealmente cada captura es de 2 metros porque esta es la longitud de la zona de estudio.

En promedio la distancia de captura entre imágenes oscila entre 1.9 y 2.1 metros, por lo que existe un desfase de unos 10 centímetros aproximadamente, lo que provoca que en ocasiones las imágenes se solapen y en otros casos haya una pequeña zona que no se trate. Este error proviene principalmente por la imprecisión del sistema GPS, descartando problemas derivados de los tiempos de proceso, ya que éstos se han visto que son suficientes para las tareas a realizar en función de las velocidades del vehículo.

6.5 Sistema experto desarrollado para el proyecto RHEA

Uno de los requisitos del proyecto RHEA es que los elementos que forman parte del prototipo han de tener un grado de protección IP65, es decir, protección contra el impacto de objetos sólidos y la penetración de agua. Además en este caso, es especialmente importante disponer de una protección contra el polvo de manera que bajo ninguna circunstancia éste entre en contacto con los elementos que forman parte del sistema de visión, y una protección contra el agua, de forma que impida la entrada de agua arrojada a chorro.

Dado que tanto la IMU como la cámara no disponen de tal protección, ambos dispositivos han sido montados dentro de una carcasa con protección IP66, la cual supera los requisitos del mencionado proyecto ya que soporta chorros más potentes de agua. Esta carcasa además está equipada con un dispositivo que controla la temperatura dentro, con el fin de que tanto la cámara como la IMU estén dentro del rango de temperatura de trabajo necesario para su correcto funcionamiento. La carcasa, figura 6.14, cuenta con un ventilador y una resistencia que actúa a modo de calefactor impidiendo que el cristal se empañe en caso de temperaturas bajas. Ambos elementos necesitan 12 V de alimentación y están regulados mediante un termostato. Su construcción es en aluminio e incluye un pequeño parasol en la parte frontal para evitar que incida directamente la luz solar, todo ello con un peso aproximado de 1000 g. Las dimensiones útiles son de aproximadamente 200 x 90 x 90 mm y exteriores de 390 x 145 x 105 mm. Tanto la cámara como la IMU se sitúan ancladas a la base para impedir desplazamientos indeseados durante el funcionamiento del sistema debido al traqueteo del tractor.

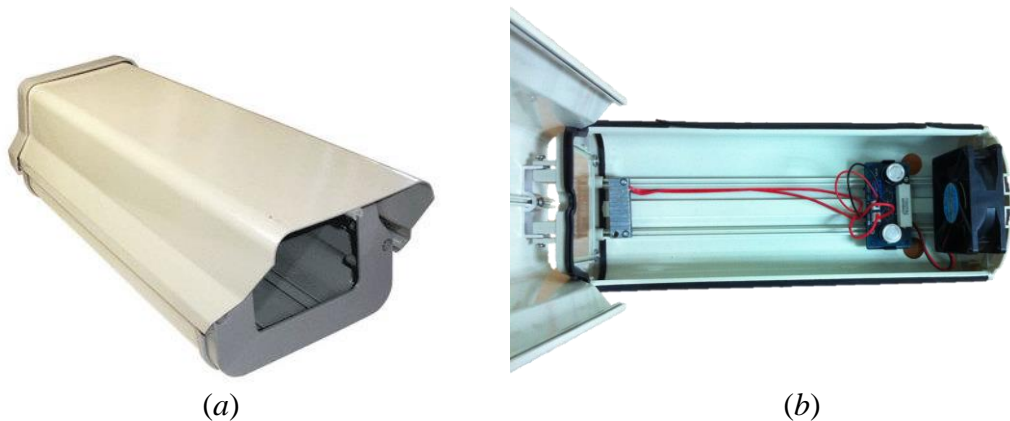


Figura 6.14 Carcasa. (a) Vista exterior. (b) Vista interior.

En la figura 6.15 se muestra el sistema de visión desarrollado en el marco del proyecto RHEA.

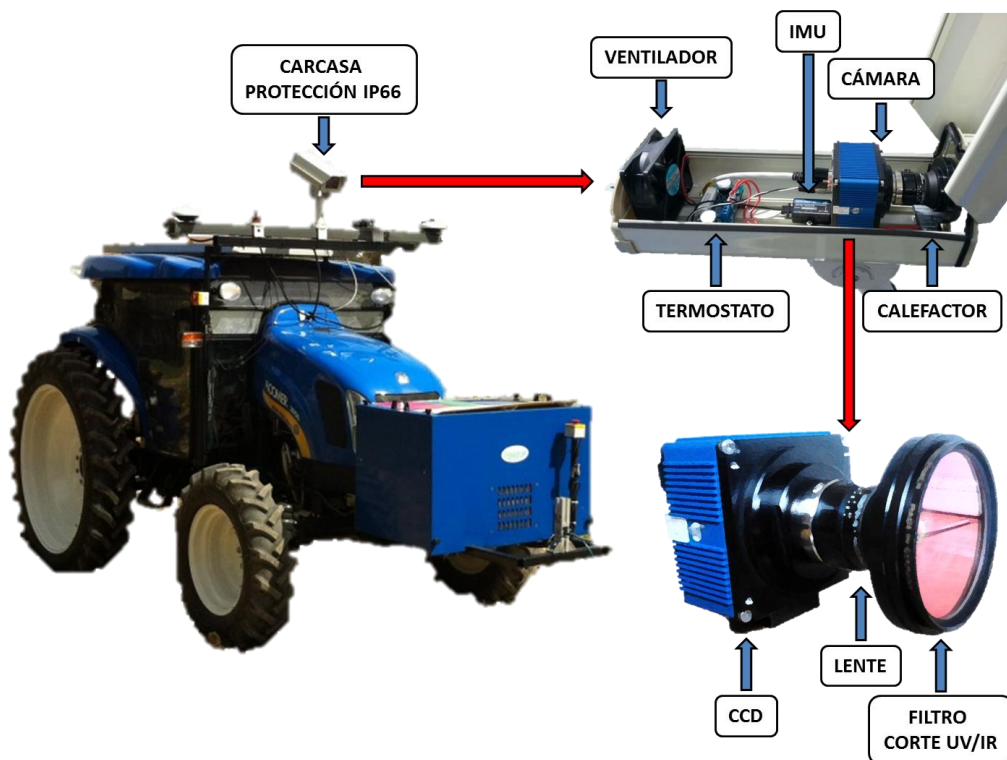


Figura 6.15 Sistema de visión instalado en el prototipo del tractor RHEA.

Éste se compone de la cámara SVS-VISTEK SVS4050CFLGEA con la lente Schneider Kreuznach Cinegon 1.9/10-0901 y el filtro Schneider UV/IR 486 además de la IMU MicroStrain 3DM-GX3-35 (AHRS) con GPS, descritos en las secciones 6.2.1, 6.2.2 y 6.2.3 respectivamente. Estos elementos se encuentran dentro de la carcasa descrita anteriormente y tanto la cámara como la IMU están situadas próximas entre sí para que los valores que suministra la IMU representen de un modo preciso la posición de la cámara.

Como plataforma de procesamiento se ha elegido la cRIO-9082 tanto por su robustez, en cuanto a diseño, como por la capacidad de realizar procesamiento a tiempo real mediante DLLs programadas en lenguaje C y embebidas en LabVIEW a través de un módulo específico. Dado que la cRIO está situada dentro de la cabina del tractor, no es necesario el uso de carcasas protectoras para este componente.

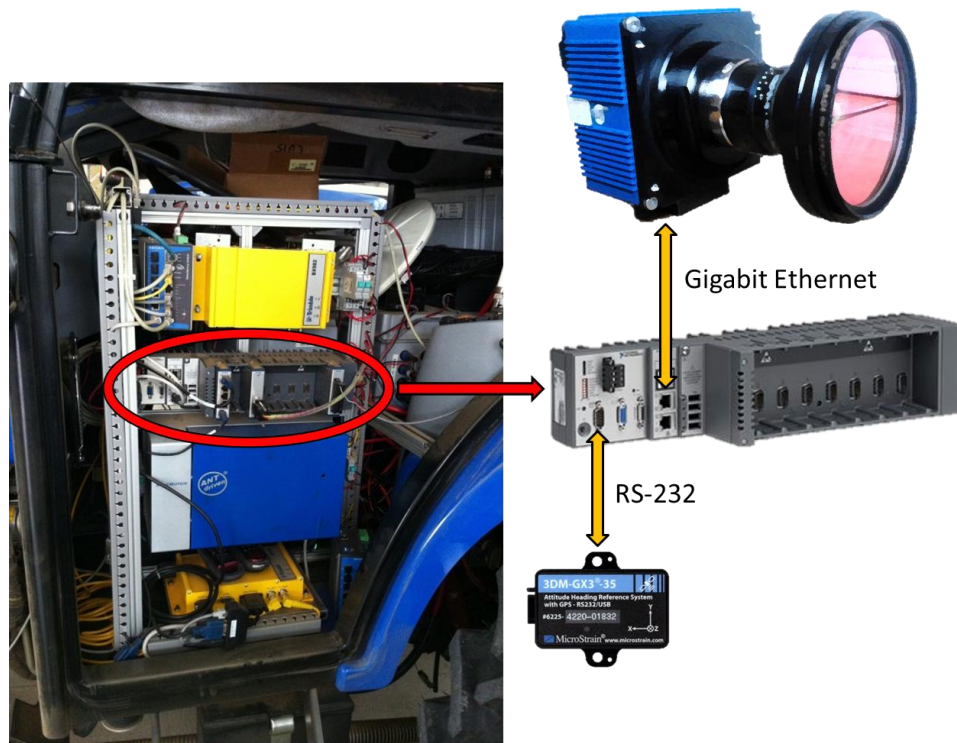


Figura 6.16 Esquema de conexión de la plataforma de procesamiento cRIO-9082 con la IMU y la cámara que forman parte del sistema de visión.

La cRIO se comunica con la cámara mediante una conexión Gigabit Ethernet (GigE) a través de una interfaz RJ-45 y con la IMU a través de un puerto RS-232, figura 6.16. La cRIO está a su vez conectada a través de otra conexión GigE a un conmutador (*switch*) para realizar el control de todo el sistema del tractor, por lo que se encarga de enviar las órdenes procedentes del sistema de visión a las partes implicadas, tales como al apero para realizar el tratamiento selectivo de las malas hierbas así como al sistema encargado de realizar el guiado.

El desarrollo de este prototipo se ha finalizado satisfactoriamente, por lo que se trata de un sistema real y completamente funcional. El diseño de la aplicación que va embebida en la DLL, tal y como ha quedado demostrado por las pruebas llevadas a cabo en el presente capítulo es suficientemente flexible y general, permitiendo cambiar tanto la cámara como la plataforma de procesamiento. La figura 6.17 muestra la interfaz de

configuración diseñada para el control del sistema de visión, que contiene todos los parámetros necesarios para la ejecución de los algoritmos y métodos diseñados e implementados.

The screenshot displays a configuration window titled 'Rutas del sistema' at the top. Below this, there are two sections: 'Ruta configuración del sistema de visión' with a text box containing 'C:\JM\configuration.ini', and 'Ruta DLL procesamiento' with a text box containing 'C:\JM\Vision.dll'. The main area is divided into four panels. The top-left panel, 'Segmentación de la imagen', contains 'Binarización' with a dropdown menu set to 'Agrupamiento/Clustering' and 'Apertura' with a numeric input set to '0'. The top-right panel, 'Tratamiento de la imagen', contains 'Vignetting' and 'Balance de blancos', both with dropdown menus set to 'NO'. The bottom-left panel, 'Detección de líneas de cultivo', contains 'Detección de líneas' with a dropdown set to 'Theil-Sen', 'Grados orientación' with a numeric input set to '6', 'Ángulos' with a dropdown set to 'IMU', and three numeric inputs for 'Pitch' (27), 'Roll' (0), and 'Yaw' (0). The bottom-right panel, 'Delimitación de la ROI en píxeles', contains four numeric inputs: 'Superior' (400), 'Inferior' (1200), 'Izquierda' (300), and 'Derecha' (1900).

Figura 6.17 Interfaz de configuración.

La interfaz es dinámica permitiendo cambiar los valores de los parámetros durante la realización de la misión de tratamiento. Dichos cambios se realizan en función de la meteorología o de las preferencias del usuario a la hora de utilizar distintos métodos y combinaciones de los mismos. En cuanto a la segmentación de la imagen permite elegir entre diferentes técnicas de binarización así como realizar la operación morfológica de apertura para aquellos cultivos donde la infesta sea mayor. La apertura se realiza con un elemento estructural de dimensiones $N \times N$, siendo N un número impar mayor que uno y menor que nueve. En lo referente a la detección de líneas permite alternar entre los métodos de Regresión y Theil-Sen, además de elegir entre utilizar ángulos fijos o los proporcionados por la IMU. Asimismo, a la hora de detectar la orientación del cultivo, permite elegir mayor o menor rango de valores para los ángulos, hasta un máximo de diez grados, siendo deseable mayor rango cuanta menos altura poseen las plantas. Además, permite realizar la corrección de la imagen con la técnica del balance de blancos y del *vignetting* así como delimitar la ROI sobre la cual se realizan las técnicas de procesamiento descritas en la presente memoria.

En cuanto a la interfaz de usuario, disponible y accesible durante la realización del tratamiento sobre los campos de maíz, se ha optado por mostrar las imágenes junto al procesamiento realizado, marcando con rectas de color verde las líneas de cultivo detectadas, en amarillo los márgenes aplicados a cada lado de la línea identificada, en rojo las líneas divisorias que se sitúan entre las líneas de cultivo y en azul los píxeles detectados como mala hierba.

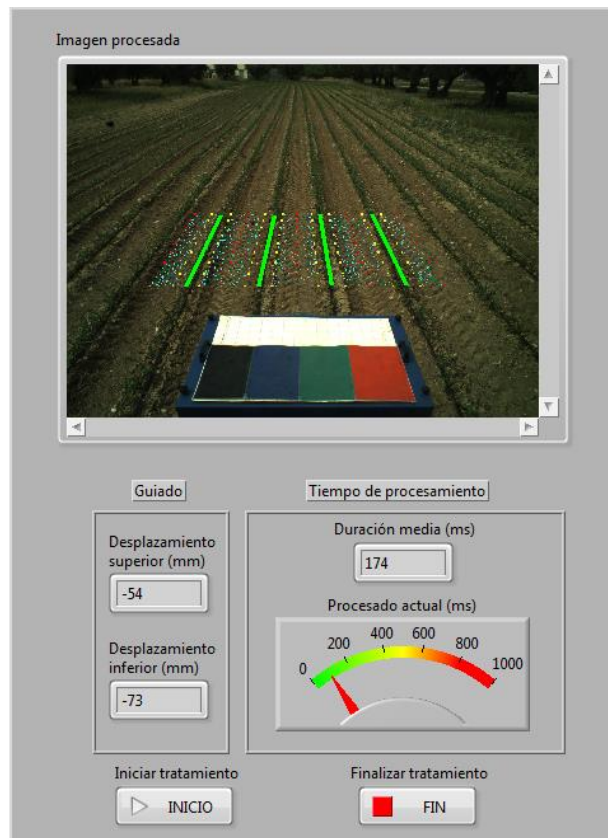


Figura 6.18 Interfaz de usuario. Vista de la aplicación durante el tratamiento.

Igualmente, como se observa en la figura 6.18, se muestra la desviación que tiene el tractor en milímetros así como el tiempo medio de procesamiento de las imágenes. El signo mostrado en el desplazamiento indica si el tractor está desplazado a la derecha (signo negativo) o a la izquierda (signo positivo).

Tal y como se muestra en la figura 6.17, es requisito indispensable que se le indique las características tanto del sistema de visión como de cultivo a tratar. Esta funcionalidad se ha implementado mediante un fichero de configuración. Esto supone asumir una configuración inicial por parte del operario final y que sólo se ha de variar cuando alguna de las partes del sistema se reemplace o bien cuando se desee trabajar

sobre otro cultivo o bajo otras condiciones, entre las que se incluye la distancia que define la ROI. La figura 6.19 muestra exactamente los valores de los parámetros de configuración por defecto utilizados para el sistema propuesto con las características de la cámara SVS-VISTEK SVS4050CFLGEA, con la lente Schneider Kreuznach Cinegon 1.9/10-0901 y con una zona de interés (ROI) situada a 3.3 metros de distancia con una longitud de dos metros sobre el campo. El factor de conversión píxel a milímetro para las características que presenta la cámara es de 1.94 mm por píxel. Las líneas de cultivo están espaciadas 75 cm entre sí y se desea identificar 4 líneas de cultivo dejando un margen de 13.5 cm a cada lado de la línea identificada (18%). Además, se quiere dividir la ROI mediante líneas horizontales espaciadas horizontalmente 25 cm de distancia, de esta manera se obtiene una matriz de porcentajes de mala hierba de 8x8 según la ecuación (5.8). Un ejemplo de la matriz de porcentajes obtenida como resultado final se muestra en la figura 5.11.

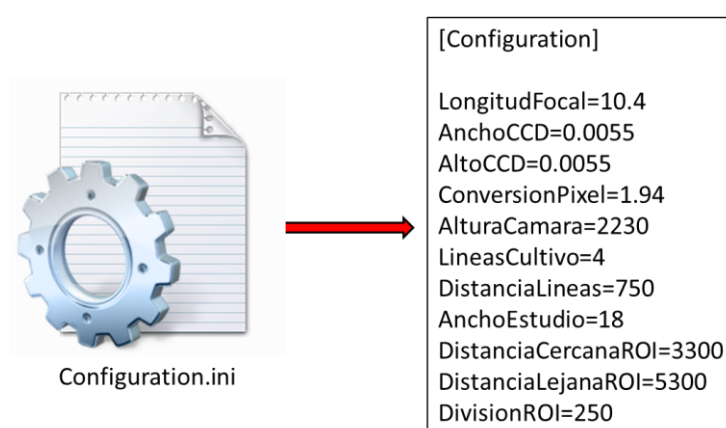


Figura 6.19 Archivo de configuración.

7

Conclusiones y trabajo futuro

7.1 Introducción

Llegados a este punto se hace necesario realizar un balance global sobre las líneas de investigación llevadas a cabo en el presente trabajo de tesis y las conclusiones obtenidas. Además, se analizan una serie de perspectivas de futuro relacionadas con los temas de investigación abordados así como sus posibilidades de ampliación.

7.2 Conclusiones

La contribución principal de este trabajo de investigación radica en el diseño de un sistema experto para AP capaz de detectar las líneas de cultivo y malas hierbas en campos de maíz además de controlar el guiado del tractor y el solapamiento de las zonas tratadas en tiempo real. El sistema está basado en dos grandes bloques. El primer bloque, encargado de obtener imágenes de calidad y su posterior segmentación para generar imágenes en blanco y negro, esto es binarias, donde las plantas aparecen en blanco y el resto en negro. El segundo bloque, donde se realiza la identificación de las líneas de cultivo y las malas hierbas, el guiado y el solapamiento de las zonas de tratamiento. El diseño de la arquitectura del sistema propuesto está dotado de un alto grado de flexibilidad que permite que cualquiera de los métodos en los dos bloques pueda sustituirse por otro, añadir nuevos métodos o eliminar alguno de los ya existentes siempre que sea necesario.

7.2.1 Segmentación de imágenes

Tal y como ha quedado puesto de manifiesto a lo largo de este trabajo, la segmentación es un proceso esencial a la hora de identificar las plantas que pertenecen al cultivo y las que son mala hierba. Por ello, la obtención de imágenes de calidad es una tarea crucial en este trabajo.

En lo referente a la obtención de imágenes y su tratamiento previo a la segmentación, se ha propuesto una estrategia para la detección de imágenes con alta intensidad lumínica que sirve como base para realizar un ajuste preciso del tiempo de exposición en aquellas cámaras con esta posibilidad. Dicha estrategia permite la obtención de imágenes de calidad con el fin de evitar la pérdida sucesiva de imágenes debida a una mala iluminación, y por lo tanto admite la adaptación del sistema a las condiciones cambiantes en exterior sin necesidad de detener la tarea de aplicación de los métodos físico-químicos de control de malas hierbas.

Por otro lado, para las cámaras que no permiten una modificación del tiempo de exposición se ha estudiado la aplicación de técnicas basadas en métodos software como son la corrección del *vignetting* y el balance de blancos. En este sentido se ha obtenido una notable mejoría con la corrección del *vignetting* y no tanto con el balance de blancos. De esto se puede inferir que una cámara que no permita el ajuste del tiempo de exposición o no disponga de un sistema óptico con lente que ajuste el iris de forma automática, no debe utilizarse en entornos de exterior no controlados ya que un funcionamiento incorrecto del sistema implicaría una errónea detección de las partes verdes y en consecuencia un tratamiento de las malas hierbas deficiente. La corrección del *vignetting* se puede aplicar en cámaras que sí permiten ajustar el tiempo de exposición, y es que este efecto, como ya se ha comentado anteriormente, también se produce por el uso de filtros de corte (ultravioleta y en mayor medida infrarrojo), por lo que si la cámara utiliza un filtro externo es posible aplicar esta estrategia sin más que ajustar los factores de corrección para el filtro.

De cara a la segmentación de las imágenes y a su binarización se han propuesto diferentes estrategias. Por un lado, la identificación de plantas y binarización basada en el modelo $YD_B D_R$ resulta ser una propuesta novedosa fundada en el sistema de codificación de televisión SECAM. Este sistema se desarrolló para ser compatible con

las imágenes en blanco y negro existentes en su momento. Debido a este requisito de compatibilidad, los estándares de color añaden a la señal básica monocroma (luminancia) una segunda señal que porta la información de color (crominancia). Así, los televisores más antiguos solamente veían la luminancia, mientras que los de color procesaban ambas señales. En base a esto, se descomponen las imágenes a color obtenidas y su estudio demuestra que es posible lograr resultados aceptables en la identificación del verde de las imágenes en un tiempo de procesamiento compatible con la implementación de sistemas en tiempo real.

Por otro lado, se ha propuesto otra nueva estrategia automática para segmentación de imágenes en campos de maíz que permite realizar un estudio pre y post-tratamiento. La principal idea es la identificación de las SVM, las cuales permiten establecer una región de separación entre clases donde se identifican las plantas enmascaradas y sin enmascarar. Este método es capaz de identificar las plantas (cultivo y malas hierbas) cuando han sido contaminadas con materiales provenientes del suelo debido al riego o a las lluvias. El enfoque propuesto también es válido para monitorizar el post-tratamiento; ya que las malas hierbas después de un tratamiento químico o térmico inician una degradación que se traduce en la pérdida del verdor existente en la etapa pre-tratamiento. El daño en el cultivo también puede ser analizado aplicando el mismo criterio de la pérdida de verdor.

7.2.2 Detección de líneas de cultivo

Para tal propósito se ha propuesto un método para la detección precisa de líneas de cultivo en campos de maíz. Una vez se obtiene una imagen correctamente binarizada se pueden identificar las plantas como píxeles blancos en la imagen binaria. De acuerdo a los parámetros intrínsecos y extrínsecos, que definen la geometría del sistema de captura de imágenes y teniendo en cuenta la proyección de perspectiva de la cámara, es posible trazar las líneas de cultivo esperadas sobre la imagen original. Considerando que los píxeles blancos en la imagen binaria se corresponden con plantas pertenecientes al cultivo y a malas hierbas, se recorre cada línea estimada y se obtiene el número de píxeles que forman parte del verde de la imagen a lo largo de dicha línea y dentro de un margen de tolerancia fijo a cada lado. Tras variar el ángulo de guiñada y detectar el ángulo más acorde al desplazamiento que sufre el tractor sobre el terreno, las líneas estimadas pueden coincidir con las líneas reales de cultivo o no. Para llevar a cabo la

detección precisa de las líneas de cultivo se aplica una corrección de las mismas mediante el estimador Theil-Sen teniendo en cuenta los píxeles anteriormente mencionados alrededor de la hipotética línea de cultivo. Esta detección se ha comprobado que es aceptable desde el punto de vista del tiempo de procesamiento para tareas agrícolas en tiempo real, robusta frente a diferentes orientaciones y estado del cultivo, y fiable, ya que mejora la precisión de la detección realizada mediante una regresión lineal simple.

7.2.3 Guiado

El guiado del tractor es de vital importancia en cultivos de maíz, ya que hay que evitar a toda costa que las ruedas del tractor pasen por encima del cultivo, ya que de otra forma lo deterioraría. La estrategia propuesta permite hacer frente a variaciones sufridas en el transcurso de la aplicación del tratamiento ya sea porque la planificación de la trayectoria difiere de la plantación real o por los errores propios del sistema GPS. Mediante la detección de las líneas y conociendo el factor de conversión de píxel a milímetro es posible corregir la trayectoria del tractor utilizando la bisectriz de las líneas centrales en el caso del estudio de un número par de líneas como es el caso propuesto, o bien utilizando la línea central en el caso de que las líneas sean impares. Los resultados obtenidos han sido satisfactorios y muestra de ello es su aplicación en el prototipo del proyecto RHEA para realizar el guiado del tractor siguiendo las líneas de maíz.

7.2.4 Solapamiento

La detección del solapamiento de las zonas a tratar es otra tarea indispensable, ya que el propósito final de la AP es la aplicación de un tratamiento específico por zonas. Por ello, se hace uso de un receptor GPS solidario al sistema de visión para tener en cuenta la posición donde se tomó la última captura y la posición relativa del tractor. Cuando el tractor ha avanzado la distancia correspondiente a la zona de interés definida y que está especificada en el archivo de configuración que contiene los parámetros específicos del sistema de visión, se realiza una nueva captura y por consiguiente se evita el solapamiento de las zonas debido a la alta tasa de adquisición de imágenes que posee la cámara. Como ocurre en el caso del guiado, tras los buenos resultados obtenidos en las diferentes pruebas llevadas a cabo, el prototipo RHEA utiliza esta técnica para el control del solapamiento de las zonas de tratamiento.

7.3 Trabajo futuro

A la vista de los resultados obtenidos y teniendo en cuenta la problemática del trabajo en entornos de exterior, la investigación llevada a cabo ha dado lugar a la aparición de nuevas líneas de interés a ser consideradas en el futuro:

- Aplicación en otros cultivos con siembra en línea: los métodos propuestos en la presente memoria se han diseñado específicamente para la detección de verdes y líneas de cultivo en campos de maíz. Una opción interesante a la par que viable es la adaptación de los métodos a otros cultivos diferentes como pueden ser los de zanahoria, tomate, patata o lechuga. Dada la flexibilidad del sistema propuesto, el énfasis en este caso habría que realizarlo en el estudio de los parámetros referentes al cultivo, ya que varía la anchura del mismo y la distancia de siembra, por lo que a priori, simplemente sería necesario modificar el archivo de configuración con los nuevos datos para abordar los nuevos escenarios.

- Aplicación en otros cultivos: la detección de malas hierbas en cultivos como el cereal donde las líneas no están definidas implica que el uso de sistemas terrestres no tengan una perspectiva general del campo donde se pueda detectar un rodal de malas hierbas estudiando las zonas de mayor densidad. Por ello, para este propósito en la actualidad existe la posibilidad de utilizar vehículos aéreos no tripulados para sobrevolar los campos y facilitar esta tarea, pero en este caso el tratamiento no se puede realizar inmediatamente después de la detección, por lo tanto, sería interesante conseguir una detección terrestre en tiempo real de estos rodales para poder realizar el tratamiento simultáneamente a la detección. Para esta detección se requiere un estudio profundo de las diferentes técnicas de segmentación, dado que las plantas y las malas hierbas son distintas, así lo es también su verdor, por lo que se propone realizar un estudio sobre las distintas etapas de crecimiento del cultivo con el fin de detectar en qué momento el verdor de las malas hierbas difiere más con el del cultivo para poder realizar una clasificación e implementar un sistema que trabaje en distintas etapas para identificar diferentes verdores realizando una clasificación apropiada. En el desarrollo de este tipo de sistemas las técnicas de aprendizaje automático son muy interesantes.

- Aplicación en otras tareas no agrícolas: saliendo del ámbito agrícola, y dada la flexibilidad del sistema propuesto, las técnicas de segmentación que en este caso intentan resaltar las partes verdes de la imagen se pueden fácilmente adaptar para la detección de distintos elementos, como pueden ser las líneas que forman los carriles en las carreteras o las vías de los trenes. Con esto, una vez destacadas las líneas en la imagen binaria, es posible aplicar la técnica de detección de líneas con los parámetros necesarios de distancia entre líneas y geometría del sistema. Así pues, de una manera sencilla se puede adaptar el sistema para un guiado de vehículos, pudiendo detectar su desplazamiento con respecto a las líneas y centrar el vehículo, además de adaptar la velocidad del mismo a la situación, por ejemplo, cuando se detecta una curva.

Parte de este trabajo ha sido llevado a cabo durante una estancia en el grupo de Visión por Computador (CV-Lab) del proyecto AutoNOMOS (2014) dirigido por el Prof. Dr. Raúl Rojas perteneciente a la Universidad Libre de Berlín. Dado que los sistemas comerciales instalados en el coche autónomo sólo son capaces de reconocer las líneas de los carriles cuando están bien definidas, mediante esta técnica se ha podido llevar a cabo una identificación tanto de las líneas de las carreteras (continuas y discontinuas) así como de los bordillos que forman parte de las aceras, permitiendo en caso de no existir líneas o de que el desgaste de las mismas sea grande, una correcta identificación que impida que el coche pueda desplazarse en exceso en el caso de que el sistema GPS falle. Además, la aproximación ha sido adaptada para detectar el ancho de la calzada y es capaz de proyectar una línea allí donde idealmente, y por la geometría del sistema, debería de encontrarse aun cuando las líneas o los bordillos no estén bien definidos o incluso si éstos aparecen ocultos por un coche u otro objeto, figura 7.1.

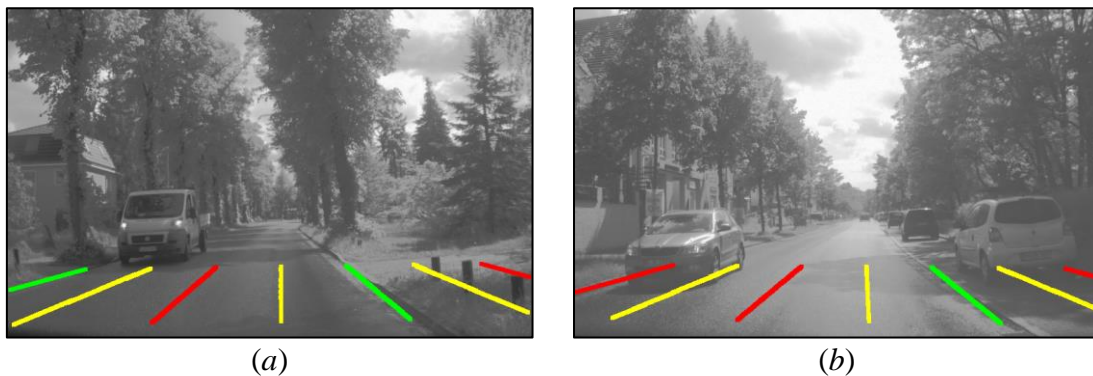


Figura 7.1 Detección de líneas en carreteras de doble sentido sin líneas pintadas, en verde las detectadas, en rojo las estimadas y en amarillo los carriles identificados (Imágenes cortesía de AutoNOMOS labs, Universidad Libre de Berlín, Alemania). (a) Carretera simple de doble sentido. (b) Carretera con estacionamiento a la derecha.

Esta técnica ha sido adaptada al sistema de visión del grupo CV-Lab, con imágenes en escala de grises y sin necesidad de alterar el sistema de visión ya integrado en el vehículo. Para llevar a cabo la detección de las líneas, el modulo encargado de segmentar la imagen ha sido modificado para realizar una extracción de bordes, figura 7.2(b), una vez se tienen los bordes extraídos, esta imagen binaria es utilizada por el módulo de detección de líneas con la configuración deseada en cuanto a líneas a detectar y configuración del sistema de visión, dando como resultado la identificación mostrada en la figura 7.2(c).

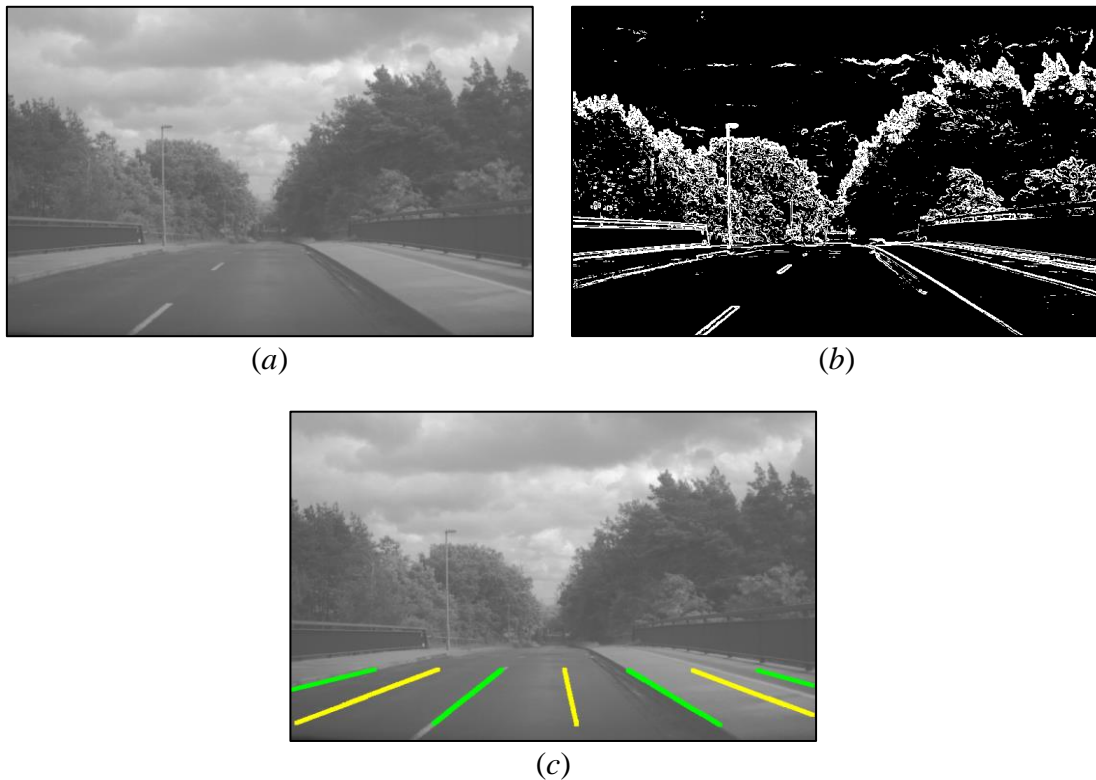


Figura 7.2 Aplicación del sistema de visión propuesto para la detección de líneas de carretera (continuas y discontinuas) y bordillos o límites de la carretera (Imágenes cortesía de AutoNOMOS labs, Universidad Libre de Berlín, Alemania). (a) Imagen original. (b) Extracción de bordes. (c) Detección de líneas y su corrección (líneas verdes pertenecientes a los bordes del carril, líneas o bordillos) e identificación de carriles (líneas amarillas).

- Detección de mala hierba en la línea de cultivo: el sistema propuesto se basa en el hecho de que las malas hierbas aparecen en rodales, y por lo tanto la aparición de las mismas próximas a la línea de cultivo implica que en las propias líneas, próximas al rodal, también exista una alta probabilidad de mala hierba. Si bien ésta es la norma en maíz y cereal, en ocasiones las malas hierbas aparecen dispersas y en los casos más dañinos, en la misma línea de cultivo. En este caso es necesario el estudio de la forma de las hojas para poder realizar una distinción, ya que en ocasiones no es posible

discernir por color. Dado que en este caso se requiere una distinción mediante la morfología de las hojas, es necesario un estudio profundo de los distintos métodos basados en regiones que fueron expuestos en el capítulo dos sección 2.4.1. Además, se requiere un nuevo estudio de la geometría del sistema, ya que con la configuración propuesta en el prototipo RHEA no es posible hacer una distinción de la morfología de las plantas debido a la pérdida de precisión por la necesidad de tener la zona de estudio relativamente alejada de la cámara. Si se quiere realizar el tratamiento simultáneamente a la detección en un escenario como el descrito, parece conveniente disponer de una vista ortogonal de las plantas para detectar la morfología con la mayor precisión posible, por lo tanto, la colocación de la cámara en una posición cenital ha de llevarse a cabo teniendo en cuenta tanto el retardo del apero a utilizar y la velocidad del tractor, como la distancia entre la cámara y el apero, ya que el procesamiento dirigido a la caracterización de la planta es más costoso y por lo tanto requiere más tiempo de procesamiento. Otra posibilidad es disminuir la velocidad del tractor, aunque dependiendo del tipo de tratamiento esto no es siempre posible. También hay que tener en cuenta que al requerir una posición cenital el ancho de la zona de estudio es muy limitado, por lo que sería necesario el uso de distintas cámaras para cubrir una mayor superficie a lo ancho del campo de cultivo. Por otra parte, la colocación de una única cámara a gran altura no es una solución viable, ya que cuanto mayor es la altura más importante es el efecto de las vibraciones y vaivenes sobre la cámara y esto provocaría desde fallos en la conexión hasta la obtención de imágenes muy borrosas y por lo tanto no aptas para la detección precisa de la forma y contorno de las hojas.

- Aplicación para dispositivos móviles: el continuo auge de los dispositivos móviles y tabletas hace que éstos cada vez sean más potentes a medida que se requiere más capacidad de procesamiento. Dado que estos dispositivos en su mayoría incluyen GPS, giróscopo y acelerómetro, es posible diseñar una versión móvil que, o bien se ejecute desde el propio dispositivo y mediante WIFI envíe la información necesaria a otro dispositivo, que tendría el operario para controlar que las tareas se llevan a cabo correctamente, o bien envíe la imagen y datos de la orientación del sistema a otro ordenador o dispositivo. Esta estrategia sería de interés para dispositivos móviles con memoria y potencia de cómputo insuficientes; en este caso el ordenador destino realizaría la ejecución y se enviarían los datos procesados de vuelta al sistema instalado en el tractor para realizar la actuación sobre el terreno.

8

English summary

8.1 Background

Weeds can cause harvest losses (both qualitative and quantitative) in global crop production. A plant is considered weed if it grows in an undesired place and its presence involve an economic damage. These plants grow in the crops and compete with them for the soil water, nutrients, space and light, reducing its performance and sometimes carrying diseases that affect the crops (Chris, 2012).

Controlling weeds in enormous land areas with a large number of hectares, it is increasingly difficult because weed plants are very competitive, with a high adaptability to adverse conditions and fast growth. Treatment by agrochemicals is the most widespread method. The problem with this practice is that some of these weeds become resistant to common herbicides, which causes the use of more expensive and powerful options applied at higher rates (GRDC, 2012). Within the agrochemical treatment, the most common practice in traditional agriculture is to spray herbicide homogeneously ignoring the spatial and temporal variability due to the high cost of unequal treatment. The large doses of herbicide and the weed resistance are a serious problem in global agriculture, which causes that uncontrolled use of agrochemicals has negative effects, most notably are the herbicide waste (financial loss) and environmental damage (contamination of the soil and groundwater).

Because of this, the use of site-specific systems becomes necessary (Gerhards and Oebel, 2006), which improves the efficiency in weed control and reduce the dose of herbicide used. The techniques aimed at optimizing the use of agricultural supplies in terms of spatial variability (differences in production in the same field, in the same season and harvesting) and temporal (changes in production in the same field in different harvest seasons) of agricultural production is known as Precision Agriculture (PA). So, the PA is to carry out the right intervention, in the right place at the right time.

The implementation of PA involves the use of tools and technologies with different levels of development, completely independent but linked from the point of view of achieving the ultimate goal (Gil, 2001). These systems can be gathered into four different technologies:

- 1) Global Position Systems (GPS), which allow to determine the position of agricultural equipment during the tasks.
- 2) Sensors, materials which can detect, process and store information about the intra-field variability.
- 3) Geographic Information Systems (GIS), tools for manage the georeferenced information.
- 4) Technologies which can control automatically the agricultural equipment, so they can modify the conditions continuously.

The latter point allows us to infer that new technologies in PA, which include machine vision systems, must be devices which can change continuously, without having to stop the task being performed. Besides, they have to be adjusted according to different adverse environmental conditions and work together with other technologies.

In machine vision systems the accuracy in weeds detection and their spatial distribution are an interesting challenge. Such vision systems have to detect weeds and make a spatial location which allows further treatment by spray-based systems or other newer techniques such as the use of burners as in the case of RHEA (2014) project, which is described below.

These vision systems can be gathered into two sets depending on whether the treatment is performed in real time or not:

- *Direct systems*: for working in real-time, these systems can act without human supervision over the areas to be treated and apply the treatment immediately after capturing the image and the weeds detection.
- *Indirect systems*: not allow to work in real-time. These systems require a subsequent data and image processing and therefore do not allow a direct action over the areas to be treated being necessary a human monitoring for the interpretation of data to identify the areas with a high infestation.

Regarding to the indirect artificial vision systems, it is noteworthy the boom that are having the drones within agriculture. A drone or UAV (Unmanned Aerial Vehicle) is an aircraft without a human pilot aboard. The flight is controlled autonomously by computers or remotely by an operator located on the ground or other vehicle. The takeoff and landing can also be performed either automatically or by using a remote control. There are different types of UAV with different sizes and configurations, many of which are planes, helicopters and multicopters (helicopters with several rotors).

While these devices are direct systems in other areas, such as the military, capable of track or locate targets, in agriculture drones are indirect systems due to they cannot act in real-time on the crop. This is largely due to the various limitations of these systems such as limited flight capacity by the fuel type or energy source, size, scope, operation area, navigation system, or the push, because powerful engines can lift heavier weight but fuel consumption increases and the height the drone can reach decreases, being impracticable to lift heavy loads such as a full herbicide spraying device. Additionally, these systems require that signals emitted and controls are safe, they are limited by the bandwidth for transmitting data and the frequency spectrum is highly saturated.

Thus, generally, the current use of vision systems on aircrafts within agriculture is focused on measuring and assessing the health of vegetation based on the measurement of radiation that plants emit or reflect, creating maps that detect photosynthetic activity discriminating among different vegetable covers, as shown in figure 1.1 (Peña et al., 2012, 2013).

For all these reasons, the indirect artificial vision systems in PA represent a great source of information but do not provide real-time performance in the field. An example of this type of system can be found in Zhang et al. (2010), who identify *Bidens Pilosa L* in wheat fields through a multimodal sensor. For this purpose, in partnership with Eaton et al. (2005) have a robot (CASPA) that can be controlled remotely or programmed for autonomous navigation using GPS and equipped with a vision system that captures images in the RGB color model and in near-infrared. As mentioned before, the drawback of this type of robot is that does not allow a direct action on the ground due to the robot not have an implement.

Unlike indirect vision systems, that cannot act in real-time, the direct systems installed onboard of ground robots or tractors allow acting in real-time on the field and are versatile to apply different techniques depending on the target in the crop.

Most of these direct artificial vision systems have been developed to detect green plants on agricultural lands by applying different techniques, outstanding those which use color images (visible spectrum) and near-infrared. An example of this system can be found in Berge et al. (2012) who using a robot equipped with a sprayer and guided by GPS, acquires images in RGB format and identifies weeds in cereal using a software tool called Weedcer developed by SINTEF (2014). This tool provides estimation data that facilitate the pixel classification on the images depending on whether they belong to the soil or to the plant.

Based on the foregoing, the desirable objectives of a direct computer vision system for agricultural tasks are:

- a) High resolution, to detect weeds and control the vehicle guidance as accurately as possible.
- b) Ability for spatial location, so that the georeferenced weeds distribution is available and the overlap of areas to be treated can be controlled avoiding studying the same area repeatedly.
- c) To be completely independent and usable with other treatment systems that use machine vision as the source of information.
- d) To act in real-time, so that the treatment is performed in-situ, without requiring study the images retrospectively and without human supervision.

8.2 Problem statement

As discussed above, nowadays the use of machine vision systems for weeds identification and its spatial location in large areas of cultivation is essential in order to simplify the tasks to be performed by the traditional operator with PA purposes.

The development of tractors and also the implements used in weeds control means that operators need very specific knowledge to perform the tasks. Therefore, the use of vision systems greatly facilitate the task carried out, and together with other systems mentioned in section 8.1 can autonomously control the tractor and also the treatments, partly alleviating time and weather restrictions.

During the early stages of maize growth, weeds compete with maize for nutrients, water and light. Therefore it is important to control weeds in order to avoid that in infested areas the maize growth is harmed, being necessary remove the weeds in early stages preserving maize plants. The ultimate aim is that the crop prevails over the weeds.

As discussed in detail in section 8.3.1, there are two main projects (RHEA and AGL2011) that require weeds identification in maize fields in order to apply site-specific treatments based on different areas and levels of infestation. It is also important to note that PA is always present from sowing to harvesting in maize fields; therefore, when talking about PA is not just about weeds identification and treatment, it is also important that the sowing is carried out appropriately. Maize is a crop whose sowing is done in straight lines, that is, seeds are placed one after the other creating crop rows whose inter-line spacing is 75 cm. Hence, weeds can appear between the crop rows or between a plant and another. Since the vision system has to work in real-time, weeds identification between plants requires a higher computational time than the identification between crop rows, concluding that if weeds exist between crop rows, there are weeds in the area next to the crop. In this process, crop rows identification as accurately as possible is a crucial task, because allow identifying the greenness between these rows to estimate the amount of weeds in the areas. Moreover, these rows are used to guide the vehicle by controlling the deviations suffered due to errors or inaccuracies in the GPS system and where the terrain, the weight of the tractor or the implement make that tractor's traction tends to deviate from its ideal trajectory.

Thus, this work is focused on the development of methods based on a direct artificial vision system for crop rows identification in maize fields and detection of vegetation densities, with the aim of applying site-specific treatments and to guide the vehicle as precisely as possible. In particular, the machine vision system is installed onboard an autonomous tractor, figure 1.2.

8.3 Motivation and goals

8.3.1 Motivation

The research developed resides in research projects carried out within the research framework of the ISCAR (2014) group, officially recognized by the Complutense University of Madrid since 2006.

The ISCAR group works on various researches, being the computer vision a particular area within robotics. Regarding machine vision systems, ISCAR group works in the projects listed below, some of them in collaboration with other national and international research groups, which are based on autonomous vehicles equipped with machine vision systems to complete different missions. One of the most important characteristics is that such vehicles operate in complex and unstructured environments, where adaptation methods to the environment and the conditions, due to the vehicles' motion, are the two major technological challenges addressed in the present research work. Here are the projects that support the research in the present thesis:

- 1) *Robot Fleets for Highly Effective Agriculture and Forestry Management* (RHEA), *THEME NMP-2009-3.4-1, Seventh Framework Programme*, (<http://www.rhea-project.eu/>) project funded by the European Community within the *7^o Framework Programme* (FP7) with duration (2010-2014). This project was very well evaluated and it was funded with 6.59M€. Nineteen research institutions and companies have participated including the Complutense University through ISCAR group, who was in charge of the computer vision in ground vehicles. The project investigates in smarter ways to protect crops using a fleet of aerial robots (drones) and terrestrial (autonomous tractors) interconnected with advanced mapping and remote sensing capabilities that enable real-time treatments and precise weeds control by spraying or by mechanical and thermal devices, or spraying trees because each treatment is

associated with a particular crop, being cereal, maize and olive respectively. For each crop there are autonomous tractors, designed as prototypes, acting together as a fleet of robots.

- 2) *Sistema de inspección terrestre en vehículos autónomos y su aplicación efectiva a la detección de malas hierbas y su control localizado* (ground inspection systems in autonomous vehicles and its effective application for weeds detection and its control) (AGL2011-30442-C02-02), belonging to the R+D+i National Plan (01-01-2012 to 31-12-2014). This project is coordinated in collaboration with Robotic and Automatic Center (CAR) -mixed center with CSIC-UPM-. The ultimate goal of the project is the spatial location of weeds patches in cereal and maize fields in order to apply site-specific treatments with agrochemicals reducing environmental pollution and production costs.
- 3) *System for Surveillance, search and rescue in the sea by means of collaboration of autonomous marine and air vehicles* (DPI2009-14552-C02-01), DPI (2009), extended to end of 2013. This project includes vision-based tasks to localize ships and castaways in the sea. These tasks required the development of methods and procedures, based on computer vision techniques in unstructured environments.

These projects propose real technological challenges in the specified areas, where computer vision involves its own complex problem, which is the objective of the proposed research in order to make vehicles as efficiently as possible without lose autonomy.

Vision systems in unstructured environments, where autonomous vehicles have to move, are characterized by changing conditions. Such is the case with the natural light, because it varies along different days, but also in the same day and in close moments can vary strongly. Also there are other problems due to vision systems are installed on mobile vehicles, where the movement (uneven terrain) or the effects produced by the engine vibration can affect significantly to the vision system and therefore to the image quality, avoiding identifying the elements of interest. For this reason, the research provides vision methods robust enough to achieve satisfactory results in the above scenarios.

As mentioned before, in RHEA and AGL2011, the problem lies in the crop rows identification in agricultural images, specifically in maize fields, with the aim of distinguishing between what is crop against weeds. The ultimate goal is to locate weeds patches to apply the treatment with mechanical, agrochemical or thermal devices in order to remove them. In this context, the correct location of crop rows and weeds is the key to success for the weeds treatment and its elimination, being also important for the tractor guidance in the field.

8.3.2 Goals

The goals in this research are listed below. All are based on the above projects and are derived from the problems arisen from the application into RHEA and AGL2011:

- a)* Study and analysis of *recognition* techniques based on Computer Vision, which aims at identifying the elements and textured in image from environments where autonomous vehicles act, outstanding the segmentation of plants (crop and weeds) and crop row identification.
- b)* Study and analysis of *spatial localization* techniques of interesting elements for autonomous navigation and its actuation. After recognition, it is necessary locate them in the space and relate them with the reference vision system.
- c)* Investigate the real problems arising from the adverse conditions of robot navigation (vibration, uneven terrain, weather) and natural light, because all affect to the image processing.
- d)* Provide solutions, improvements and changes in the methods investigated to adapt them to real-time processing.
- e)* Analysis and validation of results of the proposed solutions.

8.4 Revision of methods

Since the beginning of Robotics and Artificial Intelligence, humans have wanted to equip machines with human beings capabilities. Numerous researches try to equip visual sense into machines, because a machine with visual perception can tackle problems that remain unsolved without it. Most of the machine vision systems use other mechanisms such as range or proximity sensors (Fu, 1988) to complete the process making a multisensory integration (Lopez-Orozco, 1999). Although this research is based on the crop rows and weeds identification in maize fields applying only computer vision techniques, as mentioned in the previous section, in the results, an Inertial Measurement Unit is used to achieve the maximum accuracy as possible during detection. It is necessary to know the exact position of the vision system in relation to a virtual absolute coordinates system in order to carry out the identification process.

Since mid-60s, image processing has increased for different purposes, but they all have common characteristics such as to enhance and extract the information for its later interpretation. Therefore, Pajares and Cruz (2007) describe the artificial computer vision as the ability of a machine to see the world around it, deducing the structure and properties of the three-dimensional world from one or more two-dimensional images.

In conclusion, any process based on computer vision system implies the definition of a vision system based on one or more cameras with their optical systems capable to capture images for further processing, in this case applying artificial intelligence techniques. These two aspects are addressed in this section focused on the developed research, which is the crop rows and weeds detection for site-specific treatment and for guidance of the tractor as accurate as possible along the maize field.

Computer vision methods applied to the agriculture tasks described in this research can be considered into two categories, namely:

a) Image segmentation based on the following approaches: visible spectral-indexes (Woebbecke, 1995; Ribeiro et al., 2005; Guijarro et al., 2011), thresholding (Otsu, 1979; Kirk et al., 2009; Burgos-Artizzu et al., 2011; Meyer and Camargo-Neto, 2008) and learning methods (Meyer et al., 2004; Zheng et al., 2009, 2010).

b) Lines detection that apply the following strategies: exploration of horizontal strips (Søgaard and Olsen, 2003; Woebbecke et al., 1995; Sainz-Costa et al., 2011; Ribeiro et al., 2005; Hague et al., 2006), Hough transform (Hough, 1962; Slaughter et al., 2008; Tellaache et al., 2008*a, b*, 2011; Jones et al., 2009*a, b*; Gée et al., 2008; Rovira-Más et al., 2005), looking for the vanishing point (Pla et al., 1997; Romeo et al., 2012*a*), stereovision (Kise et al., 2005; Kise and Zhang, 2008; Rovira-Mas et al., 2008), blob analysis (Fontaine and Crowe, 2006), accumulation of green plants (Olsen, 1995), frequency analysis (Vioix et al., 2004; Bossu et al., 2006, 2009), linear regression (Guerrero et al., 2013; Montalvo et al., 2012; Billingsley and Schoenfisch, 1997; Søgaard and Olsen, 2003).

8.5 Support Vector Machines for masked and unmasked plants identification

All methods mentioned in section 8.4.1, are focused on greenness identification under the assumption that the green spectral component prevails over the red one. In this problem is not always the case, justifying the proposed strategy. Based on the considerations concerning the different methods, the proposed strategy tries to identify plants with the green spectral component masked and unmasked. Two main stages are proposed. Firstly the Otsu's thresholding approach is applied and in the second one SVM is the chosen strategy. As reported in Meyer and Camargo-Neto (2008) the advantage of using color indices is that they accentuate a particular color, which is of interest. The images analyzed contain two main dominant spectral signatures, green for plants and red for soil. For this reason greenness indices are used for plants identification and redness for soil identification. With such purpose, inspired on Guijarro et al. (2011) a combined strategy is applied because of its performance as compared against individual indices but modified as follows. In Guijarro et al. (2011) ExG, ExGR, CIVE and VEG are the four indexes combined. In the images analyzed in this work, ExGR causes that shadows generated by the maize plants are erroneously identified as green plants; this is an undesired effect in image segmentation because the computation of weeds pressure is affected. Therefore, ExGR is excluded from the combination. NDI is also discarded because many pixels belonging to the soil are classified as green ones. On the other hand, as reported in Guijarro et al. (2011) and Burgos-Artizzu et al. (2011), the automatic thresholding based on Otsu's method tends

to produce infra-segmentation, this means that only those pixels that do not offer any doubt about its belonging to plants (weeds or crop) are identified as such. This is exactly the situation offered by unmasked plants, which are to be identified at the first stage. Therefore, it is certain that pixels classified based on Otsu belong to green plants, i.e. unmasked. At the second stage SVMs are applied on those pixels that have not been identified as unmasked plants, i.e. pixels belonging to masked plants, soil and other materials. Masked plants are finally detected by identifying support vectors. Fortunately, these support vectors identify pixels belonging to masked plants or plants that have begun a drying process. Because of the learning nature derived from SVM, this strategy requires two phases: learning and decision.

8.5.1 Learning phase based on Support Vector Machines

The learning phase consists of different processes outlined in the graphic displayed in figure 4.17. According to the scheme there are three main procedures: a) Greenness identification followed by thresholding through Otsu's method to obtain a binary image, which determines two classes; b) Identification of support vectors associated to each class, based on SVM; c) Computation of average values for each set of support vectors, which establish the limits for class separation between masked and unmasked plants.

a) *Greenness identification and binarization*: given an original input image in the RGB color space, the normalization scheme of equation (3.7) is applied, which is usually applied in agronomic image segmentation (G  e et al., 2008), where R, G and B are the normalized RGB coordinates ranging from 0 to 1 obtained by equation (3.8) with $R_{\max} = G_{\max} = B_{\max} = 255$ for the 24-bit color images.

Vegetation indices to be combined are computed as equations (3.9), (3.10) and (3.11). Based on Guijarro et al. (2011) the above three indices are combined to obtain the resulting value COM as equation (4.17) where w_{ExG} , w_{CIVE} and w_{VEG} are weights for each index, representing the relative importance of each index. Guijarro et al. (2011) provide the values for the four weights participating in the combination, because in this case the *ExGR* index is excluded, the weighting values for the three indices combined are proportionally obtained, i.e. $w_{ExG} = 0.36$, $w_{CIVE} = 0.47$ and $w_{VEG} = 0.17$.

The resulting combined image COM, is linearly mapped to range in $[0, 1]$, after which, it is thresholded by applying the Otsu's method, obtaining a binary image, where white pixels identify plants in the original image, with clear spectral RGB components associated to unmasked plants. On the contrary, black pixels identify those pixels in the original image belonging to masked plants, soil and other materials present in the field.

b) *Identification of support vectors*: based on the binary image, a partition with two classes is built. Class 1 (C_1) contains those pixels coming from the original image which have been labeled as white pixels in the binary image. For any white pixel located at a known spatial position in the binary image, three spectral component values from the original image are obtained, i.e. $x = (R, G, B) \in C_1$. Similarly Class 2 (C_2) contains pixels labeled as black in the binary image. Therefore, C_1 and C_2 can be considered as two subsets containing tri-dimensional patterns belonging to two different classes. The number of total patterns is n , i.e. $n = \text{card}(C_1) + \text{card}(C_2)$. The output of the SVM system is $y \in \{+1, -1\}$, where $+1$ and -1 are associated to classes C_1 and C_2 respectively. The training set is built as (x_k, y_k) , $k = 1, \dots, n$; where each x_k belongs either to class C_1 or C_2 and y_k denotes the label it belongs to. The goal of the training process based on SVM is to find a decision function separating tridimensional data into both classes. As displayed in figure 4.18, pattern samples x_k , mapped over the 3-dimensional space appear overlapped.

Mapping input vectors into a high dimensional space through non-linear transformation functions, SVM has the ability of finding the decision function even with overlapped data (Cherkassky and Mulier, 1998; Vapnik, 2000). The SVM decision function has the general form showed in equation (4.18), where x is a generic tri-dimensional vector whose components are the three spectral R , G , B values in the original image; b is a bias constant value to be determined. Equation (4.18) defines a decision function $f(x)$ as a linear combination of kernels H , centered at each data point. A discussion about the choice of the kernel used in the proposed approach is given in section 8.7.1. Parameters α_k , $k = 1, \dots, n$ in (4.18) are the solution for the quadratic optimization problem of equation (4.19), under the constraints showed in

equation (4.20) where H is the inner product kernel introduced above and C is a regularization parameter.

In Cherkassky and Mulier (1998) is reported that there is not well-developed theory on how to select the best C , in several applications, involving SVMs, is set to a large fixed constant value, such as 2000, which is the one used in this work. Data patterns x_k associated to with the nonzero α_k are called support vectors, renamed as s_k , which are the main objective of this work, instead of the decision function as it is usual, in general. Support vectors are the most informative data points. The minimal distance to the separating hyperplane from the data point is the margin τ (Cherkassky and Mulier, 1998). A separating hyperplane is optimal if the margin is maximum. The distance between the separating hyperplane and a given pattern x_k is $y_k |f(x_k)|/\|w\|$ where w is a perpendicular vector to the hyperplane, given by equation (4.21).

The problem of finding the optimal hyperplane is that of finding the w that maximizes τ , equation (4.22), where $\|\cdot\|$ is a norm, the Euclidean is chosen in these experiments. There are infinite solutions that differ only in scaling of w . To limit solutions, fix the scale on the product of τ and the norm w . The bias parameter b introduced in equation (4.18) is computed as equation (4.23), where s_k are the s support vectors, i.e. the s data patterns associated to with the nonzero α_k , y_k is the label initially assigned to the corresponding pattern of each class $(+1, -1)$.

c) Characterization of the classes: once all support vectors are obtained, two sets of support vectors are available, S_1 and S_2 , associated to classes C_1 and C_2 respectively, where $S_1 \equiv \{s_k^1\}$ with $k = 1, \dots, n_1$ and $S_2 \equiv \{s_k^2\}$ with $k = 1, \dots, n_2$; n_1 and n_2 represent the number of support vectors belonging to each class. It is well known, under the SVM framework, that support vectors represent the most significant patterns in the class they belong to. Also that they fall just in the margin of separation between classes, i.e. close to the hyperplane. Support vectors in S_1 are close to patterns in class C_2 and vice-versa. This means that support vectors establish limits between classes. So, support vectors in S_2 are exactly the closest patterns to class C_1 . Because class C_1 represents plants with a degree of greenness, support vectors in S_2 are those patterns with the highest degree of greenness in class C_2 . A priori they were assigned to a class different to the one

containing green plants, but now they are identified as patterns with a certain degree of greenness because of their similarity to patterns in class C_1 , under this assumption support vectors in class C_2 are identified as masked plants or plants affected by the treatment, which have started the dying process. All support vectors in S_2 are characterized by computing two statistical measures, mean $\bar{x}_m = (\bar{R}_m, \bar{G}_m, \bar{B}_m)$ and standard deviation $\sigma_m = (\sigma_{mR}, \sigma_{mG}, \sigma_{mB})$. Similarly for support vectors in C_1 , mean $\bar{x}_u = (\bar{R}_u, \bar{G}_u, \bar{B}_u)$ and standard deviation $\sigma_u = (\sigma_{uR}, \sigma_{uG}, \sigma_{uB})$ where sub-indices m and u denote masked and unmasked plants respectively. Means and standard deviations are the parameters obtained during the learning phase, figure 4.17.

8.5.2 Decision phase

Once both classes are characterized by their corresponding mean and standard deviations, they are used for a decision making process.

Given a new image containing masked and unmasked plants, the goal is to identify pixels belonging to each kind of plants. Because plants are characterized by the green spectral component, the percentage of green from \bar{x}_m and \bar{x}_u and also the relative deviation with respect the green spectral components are computed by equations (4.24) and (4.25). Finally, given a pixel with its corresponding spectral components, $x = (R, G, B)$, coming from the new image, the decision rules are established in equation (4.26).

These rules establish limits of greenness. Rule 1 defines a region of greenness between support vectors S_1 and S_2 which allows identifying pixels belonging to masked plants or plants affected by the treatment. Rule 2 defines a region with high green spectral values, i.e. unmasked plants.

8.6 Expert system: crop rows and weeds detection

Machine vision systems onboard robots are being increasingly used for site-specific treatments in agriculture. With such arrangements, the robot navigates and acts over a site-specific area of a larger farm (Davies et al., 1998), where the vision systems can supply abundant information.

An important issue related to the application of machine vision methods is that concerning the crop row and weed detection, which has attracted numerous studies in this area (Onyango and Marchant, 2003; Tellaeche et al. 2008*a, b*; Burgos-Artizzu et al. 2009; Sainz-Costa et al., 2011; López-Granados, 2011; Montalvo et al., 2012; Guerrero et al., 2012). The goal is to eliminate weeds to favor the growth of crops.

The vision system consists of a CCD-based calibrated camera with known intrinsic parameters, i.e. focal length, lens distortion, image center and CCD sensor sizes and pixel resolutions. The camera is located in front of the robot, inclined with a tilt angle (pitch) and at a high from the ground. Yaw and roll angles are also known. This allows determining the rotation and translation matrices defining the extrinsic parameters. Thus, areas in the field can be identified onto the image plane. This means that given an element in the field, with its spatial location, its relative positioning on the image can be determined.

The vehicle navigates on a real terrain presenting irregularities and roughness. This produces vibrations and also swinging mainly in the pitch and roll angles. The yaw angle is assumed to be correct because otherwise the robot navigates erroneously out of the crop rows. Moreover, the spacing of crop rows in the field is also known. Because of the above, most often, the mapped expected crop rows in the image do not match with the real ones, this inaccurate estimation impedes the application of correct site-specific treatments. On the other hand the discrimination of crops and weeds in the image is a very difficult task because their Red, Green and Blue spectral components display similar values. This means that no discrimination is possible between crops and weeds based on the spectral signatures. Thus, the best option is to locate the crop rows in the image with the maximum accuracy as possible. Indeed, if the crop rows are well located, those pixels along and around the detected line as crops can be accurately identify and the remainder, which are moved away, can be considered as weeds. To achieve this goal, an automatic expert system is proposed, which exploits the human knowledge, with two main modules based on image processing techniques, as described later.

8.6.1 Expert system architecture

The system architecture is inspired on the human expert knowledge about the specific application and also considering the requirements that must be fulfilled. Astrand (2005) and Slaughter et al. (2005) propose a list of requirements for guidance systems that can be also considered for crop row detection, which in essence is a similar problem. Knowledge and requirements are mapped as follows to build the architecture of the proposed automatic expert system for the accuracy crop row detection based on images.

- a) Both, crop and weeds, display similar color spectral components and during the treatment their growth stages are similar, i.e. with similar height in the plants.
- b) Crop rows are accumulations of green plants following specific alignments oriented to the vanishing point. Crops are sown, not manually planted, and the inter-line distances in the field are known.
- c) Weeds appear on isolated or overlapped patches with respect crops with irregular distributions.
- d) Camera system geometry is known, i.e. the intrinsic and extrinsic parameters.
- e) The tractor navigates on uneven terrains with perhaps abundant irregularities.
- f) Crop rows must be located with the most accuracy as possible, regardless the distribution of weeds patches around crop and also considering that crop plants could miss along crop lines, as a common situation.
- g) The system must work on real-time. This represents a trade-off between the speed of the tractor and the computational cost.

Based on this knowledge and requirements and also considering advantages and shortcomings of the different crop row detection methods, the automatic expert system is designed consisting of two main modules: image segmentation and crop rows estimation. Figure 5.2 displays schematically these two modules with the corresponding processes.

The proposed architecture design has a high degree of flexibility that allows any of the methods in the two modules can be replaced by others, add new methods or remove any of the existing ones. This flexibility is in line with the design proposals described in previous chapters.

8.6.2 Image segmentation

Image segmentation is focused on the separation of green plants (crops and weeds) from the rest (soil, stones and others). According to point a) in the list of knowledge and requirements above, the best option to identify weeds and crop is the application of vegetation indices instead of methods based on height discrimination. Vegetation indices are well tested methods, Guijarro et al. (2011) propose a combination of vegetation indices, which is the one chosen in this case because its performance in maize fields:

a) Combination of vegetation indices

According to Guijarro et al. (2011) the indexes explained in chapter three section 3.3 are combined to obtain the resulting value COM as equation (3.13), where $w_{ExG} = 0.25$, $w_{ExGR} = 0.30$, $w_{CIVE} = 0.33$ and $w_{VEG} = 0.12$ are the weights for each index, representing their relative relevance in the combination. The resulting combined image COM, is linearly mapped to range into the interval [0, 1].

b) Greenness reinforcement

Romeo et al. (2012a) propose a fuzzy clustering strategy where the cluster containing pixels belonging to green plants has been analyzed. Clusters contain pixels with the three spectral components in the RGB model as features. Obviously, and as expected, the green spectral component is dominant. On average this component in the cluster center for green plants represents values above the 36% with respect the other two components. Exploiting this knowledge and applying the trivial reasoning that pixels coming from plants should have their green component dominant, the greenness in COM is accentuated by multiplying their values by g in equation (3.7), i.e. a new greenness is obtained as: $GA = COM * g$. The multiplication is carried out pixel by pixel and GA is linearly mapped to range in [0, 1]. Because g represents the percentage of the green component, the result obtained represents the emphasis in the greenness.

c) Umbralization

Given the transformed image GA, the next step is its binarization for posterior processing. An easy threshold based on the mean gray level of the image (histogram) has been implemented in Gée et al. (2008) where the living plant material (crop or weed) appears as white spots and the rest (i.e. soil surface, stones, shadows) as black. Also in Guijarro et al. (2011) the well-known Otsu's method, traditionally applied for binarization, has been applied. More complex approaches have been also applied such as the one used in Bossu et al. (2009), based on the K-Means clustering method. The Otsu's method is chosen for its well-known performance as reported in Meyer and Camargo-Neto (2008) and also based on the study of Sezgin and Sankur (2004) where its performance has been tested in images where the number of pixels in both parts of the image histogram that Otsu's produces is close to each other.

Figure 5.3(a) displays an original image in the RGB color space of a maize crop field. The color space transformation by applying GA is displayed in Figure 5.3(b). Figure 5.3(c) displays the image transformation from image in figure 5.3(b) by applying the Otsu's method.

8.6.3 Crop rows estimation

This module is intended to apply the knowledge embedded in points b) to e), section 8.6.1, at the same time it provides specific solutions for the requirements f) and g).

The tractor navigates on uneven terrains with perhaps abundant irregularities. The knowledge of extrinsic parameters of the vision system does not suffice because the camera is continuously involved in a permanent swinging. A customization of the Theil-Sen regression estimation approach is proposed, because of its well-tested performance in statistics.

8.6.3.1 Tracing expected crop rows

Because the crop rows arrangement are known in the field and also the extrinsic and intrinsic camera system parameters, the expected crop row locations in the image can be estimated and mapped as known lines onto the image (Montalvo et al., 2011;

Fu et al., 1987; Hartley and Zisserman, 2006). Under the assumption of ideal system geometry the expected lines should match and overlap the imaged real crop rows. Nevertheless, due to uneven terrains and errors in the crop row alignment during the sowing, this often does not occur, figure 5.5(b).

Therefore, under the above consideration, two cases can appear with respect to the expected and the imaged real crop lines: a) they match; b) they do not match. In the first case, the detection method needs to verify this matching. In the second case, a line location correction must be applied until the real crop row is located. Under this approach the system geometry through the intrinsic and extrinsic parameters guides the crop row detection process.

Now the question is: how is it possible to verify that the expected lines match or not with the real crop ones? Because white pixels representing green plants are available in the binary image, a straight line can be adjusted for specific pixel alignments that are expected to identify crop rows. This will represent the real crop line. Because the expected straight line equation and the adjusted one are available, it is possible to verify the correct or incorrect match for both lines. Thus, the effort is focused in methods for estimating the parameters defining real crop lines.

8.6.3.2 Correction of the expected crop rows: Theil-Sen estimator

An important problem to be addressed in this approach is that the method selected can cope with specific pixel alignments but also must be robust enough to avoid significant deviations caused by weeds that are not aligned and placed more or less near the main crop row alignments. This is the main issue addressed in this section.

Stewart (1999) provides a tutorial oriented toward robust parameter estimation in computer vision. Two frequently techniques used are least-median of squares (LMS) (Rousseeuw et al., 1984) and M-estimators (Hampel et al., 1986; Huber, 1981), but a huge volume of data implies that parameter estimation techniques in computer vision are heavily over constrained, even for problems where low-level feature extraction, such as edge detection, are applied. This in turn implies that parameter estimation problems in vision should be solved by least squares or, more generally, maximum likelihood estimation (MLE) techniques. Unfortunately, computer vision data are rarely drawn from a single statistical population as required for effective use of MLE.

In this approach it is necessary to estimate two parameters, defining the straight line equations associated to the corresponding crop rows, they are the slope m and the intercept b . For the linear regression approach, after several studies, it is observed that in the least squares estimator of a regression the coefficient m is vulnerable to gross errors and the associated confidence interval is, in addition, sensitive to non-normality of the parent distribution. With other measures, for example, the breakdown point (Rousseeuw and Leroy, 1987) a small number of outlying data can cause an estimate to diverge arbitrarily far from the true estimate. From the point of view of this approach, this means that few weed pixels can move the least squares fit far from the true fit, i.e. far of the real crop line. A second measure of robustness is the influence function (Hampel et al., 1986; Huber, 1981), in which the change in an estimate caused by insertion of outlying data, using a function of the distance, also causes false estimations because it should tend to zero with increasing distance to achieve robustness.

Alternative estimators for the regression coefficient, m , based on suitable rank tests are proposed by Mood and Brown (1950). They apply the estimation of both parameters m and b simultaneously using the statistical median by trial and error. Adichie (1967), proposes a more restrictive method under the assumption that the set of points to be adjusted is an absolutely continuous and symmetric distribution function with also an absolutely continuous and square integrable density function; Theil (1950) proposes a very simple estimator for m using also the statistical median; Dytham (2011) and Sen (1968) study a simple and robust estimator for m based on Kendall's (1955) tau rank correlation, a simple non-parametric test that can be used instead of normal regression. Hence, the estimator for m and b is the median of the set of slopes, where a simple slope is computed between every possible pair of pixels i and j with image coordinates (x_i, y_i) and (x_j, y_j) respectively, finally the median slope is then selected as the best estimate for m .

Based on the above considerations, the Theil-Sen estimator as proposed in Massart et al. (1997) is selected, because of its statistical efficiency and its robustness, even for low image resolutions, resulting in a promising approach in agricultural images containing crop rows.

A straight line is represented by its slope m and its intercept b as equation (5.4). Given a distribution of n pixels the goal is to adjust a straight line to such distribution.

The Theil-Sen estimator evaluates pairs of pixels i and j and computes the slope over the set of all possible pairs of such pixels, i.e. over the $n(n-1)/2$ possible combinations. This is carried out applying the equation (5.5).

The estimation of the intercept b , equation (5.6), is computed as the statistical median of the intercepts obtained with the robust slope m in (5.5). The ε parameter, set to 10^{-3} , is introduced to avoid exactly vertical lines with slope toward $\pm \infty$. Polar coordinates could be used to avoid this problem. Nevertheless, vertical lines do not appear in the real application.

As we can observe in figure 5.8, the correction of the estimated lines by Theil-Sen estimator (green lines) provides a correct orientation, making that estimated lines (red lines) match as accurately as possible with the real ones.

8.6.4 Weeds density detection

Once crop rows have been detected, the ROI is divided into rectangles or cells with the same area applying the geometry perspective. The idea is based on the fact that all green pixels identified in these cells belong to weeds, allowing obtaining the weeds density per cell by equation (5.7), where N_{pg} and N_{pt} are the number of green pixels and total number of pixels in each cell respectively.

Because crop rows have a certain width due to the leaves, it is necessary to discard a margin close to the crop row in order to avoid counting crop pixels as weeds. Also the bisector between crop rows is calculated. Figure 5.10 shows the width mentioned marked by yellow dots. In this example the width discarded is equivalent to 135 mm on the terrain at each side, so all green pixels detected between two consecutive yellow dots belong to crop, while the remainder would be identified as weeds. Figure 5.11 shows an enlarged ROI with its corresponding weeds percentage matrix.

8.7 Results

Upon presentation of fundamental techniques related to the image segmentation and crop rows and weeds identification, it is essential the analysis of the results obtained with the proposed methods, from the qualitative and quantitative points of view. Thus,

this section presents the results for the various components involved in the overall process that altogether conform the vision system that was operative within the prototype developed under the RHEA project (2014), and also AGL2011 project. The final demo in the RHEA project took place in May 21, 2014 with successful results.

8.7.1 Support Vector Machines for masked and unmasked plants identification

The images used for this study were acquired with a HPR817 digital in four different days in April/May 2007. All acquisitions were spaced by five/six days. A set of them were obtained in a pre-treatment phase after the field was watered artificially and when the field received different amounts of rainfall, figures 4.16 (a) and (b) are two representative images of this set. A second set of images were acquired in a post-treatment phase after applying a doses of herbicide, where weeds have started its decease process. Figure 4.16(c) is a representative image of this set. Due to the difference of days in the capture, they were also acquired under different illumination conditions. This circumstance does not affect the performance of the proposed process and therefore it is not required any further study with regard to lighting.

These digital images were captured under perspective projection containing only soil and plants, i.e. without panoramic sky. They were stored as 24-bit color images with resolutions of 800x600 pixels, and saved in RGB (Red, Green and Blue) color space. The images were processed with the Image Processing Toolbox from Matlab R2009a (TheMathWorks, 2012). A set of 210 images were processed.

The proposed strategy is focused on the extraction of pixels belonging to unmasked and masked plants or pixels affected by the treatment. It is based on the learning phase described in section 8.5.1. From the above set of available images 70 were used for identifying support vectors, which is the training set.

With respect the choice of kernel H in equations (4.18) and (4.19), five images were selected from the set of 70 images used for learning. 1000 pixels are manually selected, 500 labeled with $y_k = +1$ (class C_1) and 500 labeled with $y_k = -1$ (class C_2). Then, the 80% of pixels are randomly selected from each set of 500, which are used as training patterns, from which the decision function in (4.18) and the parameter b in

(4.23) are obtained. The remainder 20% of selected patterns is used for validation. So, given a pattern x it is assigned to class C_1 if $f(x) > 0$ or to class C_2 if $f(x) \leq 0$. The following three kernels have been analyzed:

- a) *Gaussian Radial Basis Functions*: $H(x, y) = \exp\{-\|x - y\|/\sigma^2\}$, σ^2 is varied from 1 to 10 in steps of 0.5 and for each σ^2 the percentage of success obtained with the validation data patterns is computed. This process is repeated 10 times. Finally, $\sigma^2 = 3.5$ is obtained with an averaged percentage of successes of 84.2%.
- b) *Polynomial*: $H(x, y) = \langle x, y \rangle^d$, with $\langle x, y \rangle$ defined as the inner product. Following the same procedure as above, but varying the parameter d from 1 to 5 in steps of 0.5, $d = 2.5$ is obtained with 79.1% of successes.
- c) *Sigmoid*: $H(x, y) = \tanh(\rho \langle x, y \rangle + \gamma)$, as before, ρ is varied from 1 to 10 in steps of 0.5 and from 1 to 5 in steps of 0.5 obtaining $\rho = 2.5$ and $\gamma = 3.0$ with 74% of successes.

Finally, the Gaussian Radial Basis function was chosen because of its best performance with $\sigma^2 = 3.5$.

Each binary image is obtained through the combined strategy, COM, from three vegetation indices followed by the Otsu's thresholding. Figures 4.19 (a) and (b) displays the COM image and the resulting binary image after applying Otsu, both obtained from the original image in figure 4.16(a), which is a representative image from the set used for learning.

Figure 4.18 displays the distribution of pixels, in the tri-dimensional RGB color space coming from the training set. Pixels labeled in green are patterns belonging to class C_1 , i.e. white pixels in the corresponding binary images. On the contrary, pixels labeled as red are patterns belonging to class C_2 , black pixels in the binary images. The number of pixels used for training was $800 \times 600 \times 70$. On average, the percentage of pattern samples belonging to classes C_1 and C_2 was respectively 35% and 75%. Figure 4.18 also displays both kinds of support vectors surrounded by two ellipses,

black points represent support vectors associated to class C_1 , i.e. belonging to unmasked plants; blue points belong to masked plants.

As mentioned before, the theoretical SVM framework strictly establishes that support vectors are those associated to with the nonzero α_k , in these experiments α_k values where absolute values are below a threshold set to 10^{-3} are discarded. On average over the total of images used for training, the number of support vectors per image was 106, with 60% and 40% associated to classes C_1 and C_2 respectively.

Table 6.4 displays values for the following three parameters (w , τ , b) computed through equations (4.21), (4.22) and (4.23) over the set of 7420 support vectors extracted from the set of 70 images.

From results in table 6.4, two important remarks are worth emphasizing. The first is that the number of support vectors obtained by the proposed SVM, which represents only the 0.022% of the total patterns processed. The second is the very small value obtained in the margin τ ; because it represents the minimal distance from the separating hyperplane from the data point, this means that data points are very close to the hyperplane and consequently classes C_1 and C_2 appear overlapped. However, the proposed strategy is able to distinguish between two types of plants, as expected.

Table 6.5 displays the parameter values resulting from the support vectors for classes C_1 and C_2 during the learning phase. They are averaged values for the green spectral components (\bar{x}_u, \bar{x}_m) , standard deviations (σ_u, σ_m) , ratios (r_{uG}, r_{mG}) , and tolerances (t_{uG}, t_{mG}) of green.

From results in table 6.5, we can see how the green spectral component in C_1 is greater than the red one. This establishes the lower limit, which is sufficient to identify unmasked plants characterized by this behavior. This does not occur in C_2 , where the red spectral component is dominant, although in a small quantity; this is due to the reflectance produced for clay elements coming from the soil, which have contaminated part of the original plants. Also they can come from plants that are started the dying process because of the treatment. We can also observe the relative small difference

between ratios r_{uG} and r_{mG} with the margins of tolerance expressed by t_{uG} and t_{mG} ; however, the SVM framework is able to distinguish between masked, unmasked or plants affected by the treatment and also between soil pixels.

Figure 4.20(a) displays the final segmented image from the original image in figure 4.16(b), where green and red labels inserted over the original image correspond to unmasked and masked plants respectively. We can see how most green labels are associated to crops and most red labels to weeds or crops close to the ground. In this last case, plants have been impregnated by materials coming from the soil, i.e. masking the greenness. In the first case, maize plants have not been impregnated, among other reasons because their height above the ground has avoided it. This circumstance also appears for some weeds where, as before, their height avoided the impregnation.

In relation to images available for post-treatment analysis, represented by the image in figure 4.20(b), we can observe from this image how those parts close to soil appear as masked and also that an important amount of weeds on the central inter-row crops appear highly affected by the treatment having started their drying process.

Additionally, in order to assess the validity of the proposed strategy 40 original images are selected randomly, which were visually analyzed by an expert to identify weeds plants and crops. The human visual observation is carried out for each image guided by the segmented image through the approach proposed in section 8.5. The expert concentrates his major effort in identifying the most troubled plants, i.e. those called masked plants. Incorrect assignments are manually marked, corrected or removed, generating a new-segmented image, considered as the ground-truth. On average, over the set of images tested, a percentage of success of 93.1% is obtained.

8.7.2 Crop rows estimation: Theil-Sen estimator

The images used for this study belongs to maize crops. They were captured with the BASLER scA1400-17FC color camera during April/May 2011 in a 1.7-ha experimental field of maize on La Poveda Research Station, Arganda del Rey, Madrid. All acquisitions were spaced by five/six days, i.e. they were obtained under different environmental lighting conditions and different growth stages in maize and weed plants. The digital images were captured under perspective projection and stored as 24-bit color

images with resolutions of 1392×1038 pixels saved in RGB (Red, Green and Blue) color space in the TIFF format. The images were processed under LabVIEW Real-Time (2012) from National Instruments, release 2011, under a cRIO-9082 1.33 GHz dual-core Intel Core i7 processor, including LX150 FPGA with Real-Time Operating System. The proposed algorithm is developed in C with MS Visual Studio and compiled as a DLL, which is embedded as an additional module in LabVIEW. A set of 240 images was processed. This equipment is intended to fulfill the real-time specifications expressed in chapter four section 8.6.1 point g).

The extrinsic and intrinsic camera parameters are: number of crops = 3, pitch angle = 20° , roll angle = 0° and yaw angle = 0° with the camera placed at a height of 1.5 meters from the ground; the focal length was 8.5 mm and the ROI is placed in front of the camera from 2 meters to 6 meters.

An illustrative result, displayed in figure 6.11, is the outperformance of the proposed Theil-Sen estimator as compared to linear regression.

The green pixels in the right crop line are the annotated pixels obtained by considering the expected crop line as given by the application geometric transformations according to the intrinsic and extrinsic parameters and the margin of tolerance set to 37.5 cm for each side around the expected crop line (50% of the distance between crop rows). As we can easily infer, pixels far away from the central ones are weeds, i.e. they appear scattered in the inter-rows. Red line in the image is estimated by applying the Theil-Sen estimator and blue line is the one estimated by linear regression. As we can easily see, the best adjusting is achieved by the Theil-Sen's method. This is because it is robust enough against pixel dispersions. On the contrary, the regression-based method is sensitive to this kind of dispersion because it is based on the computation of minimum distances and the scattered pixels exert an important attractiveness.

The performance of the Theil-Sen estimator against the linear regression is studied through a qualitative analysis, based on the human expert criterion because no ground truth images are available. By visual inspection of real crop row on the images, the expert determines the best adjustment of the estimated crop lines.

Because this approach is oriented to work in real-time applications once the quality performance is achieved, the behavior of both estimators are analyzed when they are applied to images with different resolutions, now under a quantitative analysis.

Figures 6.12 (a) and (b) represent an illustrative example representing an image of the 240 ones analyzed. The area of interest is delimited by landmarks with a wide of 2.25 m (covering three crop rows with inter-row spacing of 0.75 m) and 4 m long. This area is the one to be processed during a normal operation of the agricultural robot.

8.7.2.1 Qualitative analysis

Figure 6.12(a) displays in green those pixels representing green plants (crop and weeds). The expected crop lines, according to the system geometry based on extrinsic and intrinsic parameters, are drawn as yellow lines. A simple image inspection based on the human expert criterion allows us to infer that the expected crop lines do not match accurately with the real ones. This observation is particularly relevant in the left crop line. This situation is due to the fact that this crop row was the outer line applied by the seeder machine and its location differs from the 0.75 m of the ideal inter-line spacing in maize fields. The deviation of the expected central and right crop lines from the real ones is less marked than in the left one. It is obvious that under this situation the expected crop lines need correction. The three real crop lines contain some gaps produced by errors during the sowing or perhaps because the maize seeds have not emerged.

Figure 6.12(b) displays corrections of the yellow lines by applying both the Theil-Sen estimator (red lines) and linear regression (blue lines). Based on the human expertise, it is easy to see how the three red lines are well adjusted by Theil-Sen. Indeed, they follow the central part of the furrow. Perhaps in the central crop line a slight deviation can be appreciated, which compared against the one produced by the linear regression becomes irrelevant. The huge deviation produced by this last estimator in the right crop row is due to the presence of isolated patches (marked with the circle), probably weeds, which have been considered during the procedure of estimation. Something similar happens with respect the left crop rows, but the deviation it is less pronounced.

This is the general behavior observed in the set of 240 images analyzed, which allows us to verify the outperformance of the Theil-Sen estimator as compared to the linear regression. The margin of tolerance used was again 37.5 cm for each side. This allows us to conclude that Theil-Sen works appropriately under this type of situations which are abundant in maize fields.

In order to detect the real crop lines over the image the Hough transformation (Slaughter et al., 2008) is also applied. Despite geometric constraints considering the extrinsic and intrinsic parameters are applied; the method produces abundant peaks on the accumulator cells, making difficult the selection of the correct one that identifies a crop row. A lot of lines, with different slopes, appear for a unique crop row. This requires a careful peak thresholding selection (Jones et al., 2009*a, b*; Rovira-Más et al. 2005).

8.7.2.2 Quantitative analysis

From a point of view of quantitative analysis the average percentage of success based on the human expert criterion are compared for both linear regression and Theil-Sen and also without adjustment. Each image was visually analyzed by an expert to identify the accuracy between the adjusted line and the real crop row considered as satisfactory by the expert. The computational cost measured in processing times for these three approaches is also computed. Without adjustment (WA) is the procedure involving GA extraction and binarization based on Otsu's method, i.e. analyzing only the crop line detection obtained from a direct geometric mapping based on intrinsic and extrinsic parameters.

For linear regression and Theil-Sen, times displayed are exclusively the ones obtained for the specific adjustment, i.e. to obtain the total time they must be added to the one displayed without adjustment.

Because Theil-Sen is computational expensive but effective, the resolution of the ROI is reduced by down-sampling until to achieve values of 30% and 60% of the original image to verify the performances with and without reductions. Table 6.7 displays averaged percentages and computational times expressed in milliseconds for the set of original images available (i.e. with the 0% of reduction) and also for these images when their ROIs are reduced until the 30% and 60% respectively, in horizontal

and vertical sizes. For each reduction, the averaged percentages and times for three values (5%, 18% and 50%), which is called margin of tolerance, are displayed. This margin represents the horizontal width used to search pixels representing crop and weeds around the expected crop line obtained by applying geometrical constraints (extrinsic and intrinsic parameters), equivalents in the real terrain to 3.75 cm, 13.5 cm and 37.5 cm respectively, considering 75 cm the distance between crop rows. Under this consideration all values for WA do not vary because here no search is required.

From table 6.7 we can see that the best performance, in terms of accuracy, is achieved by the Theil-Sen approach as compared to WA and linear regression, because it obtains the best results. The best absolute performance of Theil-Sen, also in terms of accuracy, is obtained for the original ROI with a margin of 18%. This is because these images contain the maximum information and this margin covers the necessary range to capture all information coming from crops being unaffected by weeds pixels. As the reduction increases the accuracy decreases. The decreasing can be explained by the fact that the greater the reduction and tolerance, more pixels belonging to weeds are involved in the line estimation. Different combinations of margins and reduction have been tested without apparent improvements with respect to the ones displayed.

Regarding times, it is obvious that the greater the image sizes and margins of tolerance the greater are processing times. This is because a greater number of pixels need to be processed. The main drawback for Theil-Sen is its high computational cost. Assuming the vision system captures an area of 4 m long the robot will need to navigate at speeds below of $4/T$ (m/s) to gain time for image processing and actuation. T represents the processing time. So, in the worst case $T = 571$ ms (146 ms of WA + 425 ms of Theil-Sen), the corresponding speed is 25.21 km/h, excessive speed for agricultural tasks. This means that it is unnecessary reduce the ROI because the vision system ensures computation times within the normal operation speeds, which range between 3 and 6 km/h for thermal treatments, and between 6 and 8 km/h for herbicide treatments.

For the sake of clarity, figure 6.13 displays averaged processing times, in seconds, over the set of images available for the full process including greenness extraction based on GA, binarization through the Otsu's method and line adjustment based on the Theil-Sen estimator. As above, these values are obtained for the three

values of reduction (0%, 30% and 60%) and with the three margins of tolerance (5%, 18% and 50%). As expected, processing times decrease as the reduction and margin of tolerance also decrease.

In order to validate this strategy with a different system, this technique has been applied to the images provided by the SVS-VISTEK SVS4050CFLGEA (2014) camera after detecting the correct orientation. The images, with a resolution of 2336 x 1752 pixels and saved in RGB (Red, Green and Blue) color space in the BMP format, were processed in a cRIO-9082 in real-time through a DLL implemented in C and embedded within a LabVIEW 2011 module. The images were acquired in the CSIC-CAR in Arganda del Rey (Madrid) on May 2013 in a maize field, with the tractor stopped and also moving along and around the field. Because these cameras are different, they provide images with different characteristics. Besides, the configuration of the vision system is completely different, both in terms of height (2.23 m) and in relation to the angles (pitch = 22°, roll = 0° and yaw = 0°), the specifications of the ROI (from 3.3 m to 5.3 m), focal length (10.4 mm) and the number of crop rows (4 crop rows). Thus is demonstrated the flexibility of the proposed system on different configurations only changing the configuration with the necessary system parameters.

Figure 5.8 shows two representative images from this set where we can observe how Theil-Sen estimator is sufficiently robust to different system configurations and can cope with strong deviations from the initial estimation.

Table 6.8 shows the mean percentages of successful and the computational time in milliseconds for the original images available (i.e. with 0% of reduction) and also with different margins of tolerance (5%, 18% and 50%) which as in the above images are equivalent to 3.75 cm, 13.5 cm and 37.5 cm respectively.

With this camera in the worst case the time would be $T = 641$ milliseconds (162 ms of WA + 479 ms of Theil-Sen), so that the maximum speed which tractor should navigate would be 11.23 km/h, which is less than in the previous case but is also excessive for agricultural tasks.

Therefore, it is demonstrated the feasibility of this technique for agricultural applications requiring crop rows identification as accurately as possible in real-time.

8.8 Conclusions

The main contribution of this research lies in the design of an expert system for PA able to detect crop rows and weeds in maize fields, to correct deviations during guidance of the tractor and to control the overlapping of the areas to be treated. The system is based on two main modules. The first is in charge of obtaining high quality images with a posterior segmentation to obtain binary images where plants appear as white pixels and the remainder in black. The second module carries out the crop rows and weeds detection, the tractor guidance and controls the overlapping. The design of the system architecture has a high degree of flexibility which allows that any method can be easily replaced or removed, and also the addition of new methods.

8.8.1 Image segmentation

As has been shown throughout this work, the segmentation is an essential process to identifying plants belonging to crops or weeds. Therefore, the acquisition of quality images is a crucial task in this work.

Regarding to image acquisition and its processing prior to segmentation, a detection image strategy has been proposed in order to identify images with high light intensity which is the basis for an accurate exposure time adjustment in those cameras with this characteristic. This strategy allows the acquisition of high quality images in order to avoid losing of consecutive images due to a poor lighting. It also allows adapting the system to environmental changing conditions in outdoor environments without stopping the physicochemical treatments applied for weeds removal.

On the other hand, for cameras which cannot modify the exposure time, a set of software-based techniques have been studied such as *vignetting* correction and white balance, making important contributions for working in this kind of unstructured outdoor environments. For image segmentation and its binarization, different strategies have been proposed with satisfactory results.

It has also proposed a new automatic strategy for image segmentation in maize fields. The main underlying idea is the identification of support vectors, which allow establishing a region of separation between classes where masked and unmasked plants

have been assigned. The method is able to identify plants (weeds and crops) when they have been contaminated with materials coming from the soil, due to artificial irrigation or natural rainfall. The proposed approach is also valid for monitoring the post-treatment. This is based on the assumption that weeds, after chemical or mechanical treatments, must initiate a progressive degradation expressed by the loss of the greenness existing in the pre-treatment stage. The damage in the crop, when it occurs, can also be analyzed based on the same criterion because of loss of greenness.

8.8.2 Crop rows detection

A new method for accuracy crop row detection in maize fields has been proposed. The image is segmented to transform the original color image into a gray scale. Then a binarization process is carried out based on the Otsu's method. This allows identifying green plants which are white pixels in the binary image. According to the intrinsic and extrinsic camera parameters and applying perspective projection, the expected crop lines are traced over the original image. Considering that white pixels in the binary image are crop and weeds in the original one the expected crop lines are followed and the horizontal direction is explored to capture white pixels in the binary image. These pixels are the ones used for estimating a new crop line that can coincide with the expected one or not. The yaw angle is varied with the aim of detecting the best value according to the deviation incurred by the tractor, if the new line does not match with the expected one, a correction is made, but if it matches the correct location is verified. The estimation is carried out by applying Theil-Sen and also with linear regression. In terms of accuracy Theil-Sen outperforms linear regression based on qualitative and quantitative analysis being acceptable from the point of view of the processing time.

8.9 Future work

Because of the results obtained and considering the problems of work in outdoor environments, this research has revealed some interesting points that should be considered in the future:

- Use in others crops with wide rows: the proposed methods in this work are specifically designed to greenness and crop rows detection in maize fields. An interesting and feasible option is to adapt these methods to other different crops such as carrot, tomato, potato or lettuce, which are arranged as wide rows. Given the flexibility of the proposed system, only the parameters affecting the inter-lines width have to be adjusted.
- Use in non-agricultural tasks: leaving the agricultural sector and given the flexibility of the proposed system, the segmentation techniques, that in this case make green parts stand out in the image, can be adapted to detect other elements, such as the case of the road lines or the rail tracks. Once the lines are stand out, it is possible to apply the technique to detect lines with the required parameters of distance between lines and the system geometry. Thus, the system can be adapted to guide the vehicle detecting its displacement with respect to the lines and centering the vehicle. Also the speed can be adapted to different situations such as when a curve is detected.

Part of this work has been carried out during a research stay in the Computer Vision group (CV-Lab) of AutoNOMOS (2014) project led by Prof. Dr. Raúl Rojas belonging to the Free University of Berlin. Because of the commercial systems located in the autonomous car only can detect the lane markings when they are well defined, applying this techniques is possible to detect lane markings (continuous and discontinuous) and also the curbs of the sidewalks allowing the detection when lane markings are missing in order to prevent deviations due to GPS errors. Furthermore, this technique has been adapted to detect the width of the road, and also, infers the position of the lines, if these lines are not detected or if the curb is not well defined or be hidden by any other car or object, the lines can be estimated and mapped as known lines onto the image, figure 7.1.

This technique has been adapted to the new grayscale vision system without modifying its location. For carrying out this detection, the segmentation module has been modified to perform edge extraction, figure 7.2(b), once having the extracted edges, the binary image is used by the detection lines module with the new setting to detect lines and obtaining the lane identification shown in figure 7.2 (c).

- Use in other crops: crops and weeds have been identified based on the spectral RGB components through vegetation indices. Crops and weeds have different hues of green and textures that could be conveniently exploited.

- Weeds detection within crop row: the proposed system is based on the detection of weeds between the crop rows, assuming that patches in such area are extended inside the crop rows. Nevertheless, some weeds could be identified based on their shapes because of the morphology. This requires the study of different methods based on image region segmentation that were exposed in section 8.4. In addition, a new geometry system arrangement is required. The configuration proposed in the RHEA prototype makes not possible to distinguish the plants' morphology due to the loss of accuracy by the need to study an area relatively far from the camera. For this purpose an orthogonal camera configuration is required placing the camera in a zenithal arrangement. A trade-off is required to accomplish acceptable performances with respect the size of ROIs, processing times, speed of the tractor or delays on the implement actuation. Multiple cameras should be used to cover the required wide field of view. A unique camera in zenithal position should be installed at a height where vibrations could cause problems with the connections or blurred images which would be not valid for shape identification.

- Application for mobile devices: the continued rise of the mobile devices and tablets makes them increasingly more powerful as more processing is required. Because most of these devices include GPS, gyroscope and accelerometer, it is possible to design a mobile version that either run in the device itself and send the information via WiFi to another device, which the operator would have for controlling that tasks are carrying out correctly, or if the device has no the desirable memory or computing power, it would send the image and orientation data to another computer system or device that allows the execution. The processed data will be sent back to the system installed on the tractor to perform the action on the ground.

Bibliografía / Bibliography

- Ahuja, N., Rosenfeld, A. (1978). A note on the use of second-order gray-level statistics for threshold selection, *IEEE Trans. Systems Man Cybernet.* SMC-8, 895-899.
- Aiyer, A., Pyun, K., Huang, Y., O'Brien, D.B., Gray, R.M. (2005). Lloyd Clustering of Gauss mixture models for image compression and classification, *Signal Processing: Image Communication*, 20, 459-485.
- Andújar, D., Rueda-Ayala, V., Hugo Moreno, H., Rosell-Polo, J.R., Escolá, A., Valero, C., Gerhards, R., Fernández-Quintanilla, C., Dorado, J., Griepentrog, H.W. (2013). Discriminating crop, Weeds and Soil Surface with a Terrestrial LIDAR Sensor. *Sensors*, 13, 14662-14675.
- Andújar, D., Ribeiro, A., Fernández-Quintanilla, C., Dorado, J. (2011) Accuracy and feasibility of optoelectronic sensors for weed mapping in wide row crops. *Sensors*, 11, 2304–2318.
- Antti, M., Ville, H., Jorma, V. (2012). Precise online auto-focus system in high speed laser micromachining applications. *Physics Procedia*, 39, 807-813.
- Astrand, B., Baerveldt, A.J. (2005). A vision based row-following system for agricultural field machinery. *Mechatronics*, 15, 251–269.
- Atkinson, P.M., Tatnall, A.R.L. (1997). Neural Networks in Remote Sensing, *Int. Journal of Remote Sensing*, 18, 699-709.
- AutoNOMOS (2014). Página web: <http://www.autonomos.inf.fu-berlin.de/> (accedido agosto 2014)
- Axis Communications (2014). Disponible on-line: <http://www.axis.com/es/> (accedido agosto 2014).
- Ballard, D.H., Brown, C.M. (1982). *Computer Vision*. Prentice-Hall, Englewood Cliffs:NJ.
- Bartz, M.R. (1969). Optimizing a video processor for OCR, in *Proceedings, International Joint Conference on AT*, pp. 79-90.

- Benediktsson, J.A., Swain, P.H., Ersoy, O.K. (1990). Neural network approaches versus statistical methods in classification of multisource remote sensing data, *IEEE Trans. Geosci. Remote Sens.*, 28(4), 540–552.
- Berge T., Goldberg S., Kaspersen K., Netland J. (2012). Towards machine vision based site-specific weed management in cereals. *Computers and Electronics in Agriculture* 81, 79-86.
- Bhanu, B., Faugeras, O. (1982). Segmentation of images having unimodal distributions, *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-4*, 408-419.
- Billingsley, J., Schoenfisch, M. (1997). The succesful develepmnet of a vision guidance system for agriculture. *Computers and Electronics in Agriculture*, 16, 147–163.
- Birth, G.S. and McVey, G.R. (1968). Measuring the color of growing turf with a reflectance spectrophotometer. *Agronomy J.* 60:640-643.
- Bossu, J., Gée, Ch., Jones, G., Truchetet, F. (2009). Wavelet transform to discriminate between crop and weed in perspective agronomic images. *Computers and Electronics in Agriculture*, 65, 133–143.
- Bossu, J., Gée, Ch., Guillemain, J.P., Truchetet, F. (2006). Development of methods based on double Hough transform and Gabor filtering to discriminate crop and weeds in agronomic images. In *Proc. SPIE 18th annual symposium electronic imaging science and technology* (Vol. 6070), paper no. 23, San Jose, USA.
- Boukharouba, S., Rebordao, J.M., Wendel, P.L. (1985). An amplitude segmentation method based on the distribution function of an image, *Comput. Vision Graphics Image Process.* 29, 47-59.
- Burgos-Artizzu, X.P., Ribeiro, A., Guijarro, M., Pajares, G. (2011). Real-time image processing for crop/weed discrimination in maize fields. *Computers and Electronics in Agriculture*, 75, 337–346.
- Burgos-Artizzu, X.P., Ribeiro, A., Tellaeché, A., Pajares, G., Fernández-Quintanilla, C. (2010). Analysis of natural images processing for the extraction of agricultural elements. *Image and Vision Computing* (ISSN 0262-8856), vol. 28, pp. 138-149.
- Burgos-Artizzu, X.P., Ribeiro, A., Tellaeché, A., Pajares, G., Fernández-Quintanilla, C. (2009). Improving weed pressure assessment using digital images from an experience-based reasoning approach. *Computers and Electronics in Agriculture*, 65, 176–185.
- Burks, T.F., Shearer, S.A., Payne, F.A. (2000). Classification of weed species using color texture features and discriminant analysis. *Transactions of the ASAE*, 43(2): 441-448.
- Campadelli, P., Medici, D., Schettini, R. (1997). Color image segmentation using Hopfield Networks, *Image and Vision Computing*, 15, 161-166.
- Camps-Valls, G., Bruzzone, L. (2005). Kernel-based methods for hyperspectral image classification, *IEEE Trans. Geosci. Remote Sens.*, 43(6), 1351–1362.

- Camps-Valls, G., Gómez-Chova, L., Calpe, J., Soria, E., Martín, J. D., Alonso, L., Moreno, J. (2004). Robust support vector method for hyperspectral data classification and knowledge discovery, *IEEE Trans. Geoscience. Remote Sensing*, 42(7), 1530–1542.
- Chan, J.C.W., Laporte, N., Defries, R.S. (2003). Texture Classification of logged forest in tropical Africa using machine-learning algorithms, *Int. J. Remote Sensing*, 24(6), 1401-1407.
- Chang, H.C., Shih T.M., Chen, N.Z., Pu, N.W. (2009). A microscope system based on bevel-axial method auto-focus. *Optics and Lasers in Engineering*, 47, 547-551.
- Chen, Y., Jin, X., Tang, L., Che, J., Sun, Y., Chen, J. (2013). Intra-row weed recognition using plant spacing information in stereo images. American Society of Agricultural and Biological Engineers, St. Joseph, Michigan. Paper number 131592292, Kansas City, Missouri, July 21 - July 24, 2013. (doi: 10.13031/aim.20131592292).
- Chen, C.Y., Hwang, R.C., Chen, Y.J. (2010). A passive auto-focus camera control system. *Applied Soft Computing*, 10, 296-303.
- Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J. (2001). Color image segmentation: advances and prospects. *Pattern Recognition*, 34, 2259-2281.
- Cherkassky, V., Mulier, F. (1998). *Learning from data: Concepts, theory and methods*. New York: Wiley.
- Chow, C.K., Kaneko, T. (1972). Automatic boundary detection of left ventricle from cineangiograms, *Comput. Biomed. Res.* 5, 338-410.
- Chow, C.K., Kaneko, T. (1971). Boundary detection of radiographic images by a threshold method. In *Proceedings IFIP Congress 71*, page 1530-1535.
- Chris, P. (2012). Parasitic Weeds: A World Challenge. *Weed Science*, vol. 60, no. 2, pp. 269-270.
- Conesa-Muñoz, J., Ribeiro, A., Andujar, D., Fernandez-Quintanilla, C., Dorado, J. (2012) Multi-path planning based on a NSGA-II for a fleet of robots to work on agricultural tasks. *IEEE Congress on Evolutionary Computation*, pages 1-8, 10-15 June, E-ISBN: 978-1-4673-1508-1, print ISBN: 978-1-4673-1510-4, doi: 10.1109/CEC.2012.6256629
- Davies, G., Casady, W., Massey, R. (1998). *Precision agriculture: An introduction. Water Quality Focus Guide (WQ450)*, Disponible on-line: http://www.naarm.ernet.in/virtual/gis_vol2/precisionagri.pdf (accedido agosto 2014).
- Dawson-Howe, K. (2014). *A Practical Introduction to Computer Vision with OpenCV (1ª ed.)*. ISBN: 978-1-118-84845-6, 234 pages, John Wiley and Sons, United Kingdom.
- Debeir, O., van den Steen, I., Latinne, P., van Ham, P., Wolff, E. (2002). Textural and contextual land-cover classification using single and multiple classifier systems, *Photogramm. Eng. Remote Sens.*, 68(6), 597–605.

- Del-Frate, F., Pacifici, F., Schiavon, G., Solimini, C. (2007). Use of Neural Networks for Automatic Classification from High-Resolution Images, *IEEE Trans. Geoscience and Remote Sensing*, 45(4), 800- 809.
- Deravi, F., Pal, S.K. (1983). Gray level thresholding using second-order statistics, *Pattern Recognit. Lett.* 1, 417-422.
- Doyle, W. (1962). Operation useful for similarity-invariant pattern recognition, *J. Assoc. Comput. Mach.* 9, 259-267.
- Dytham, C. (2011). *Choosing and Using Statistics: A Biologist's Guide* (3rd ed.). ISBN: 978-1-4051-9839-4. p. 230, John Wiley and Sons, United Kingdom.
- Eaton, R., Katupitiya, J., Cole, A., Meyer, C. (2005). Architecture of an automated agricultural tractor: Hardware, software and control systems. In *Proceedings of the 16th World Congress of the IFAC*, Prague, Czech Republic, July 2005.
- Emmi, L., Paredes-Madrid, L., Ribeiro, A., Pajares, G., Gonzalez-de-Santos, P. (2013). Fleets of robots for precision agriculture: a simulation environment. *Industrial Robot: An international Journal*, 40(1), 41-58, ISSN 0143-991x, doi: 10.1108/01439911311294246.
- Fekete, G., Eklundh, J.O., Rosenfeld, A. (1981). Relaxation: Evaluation and applications, *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-3, 460-469.
- Fernando, S.M.X., Monro, D.M. (1982). Variable thresholding applied to angiography, in *Proceedings, 6th International Conference on Pattern Recognition*.
- Ferreiro-Armán, M., Bandeira, L.P.C., Martín-Herrero, J., Pina, P. (2007). Classifiers for vegetation and forest mapping with low resolution multispectral imagery, in: J. Martí, J.M. Benedí, A.M. Mendoça and J. Serrat (Eds.), *Pattern Recognition and Image Analysis*, Springer-Verlag, Berlin, LNCS 4477, Part I, 177-184.
- Fisher, R.A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics* 7 (2): 179–188. doi:10.1111/j.1469-1809.1936.tb02137.x. hdl:2440/15227.
- Fontaine, V., Crowe, T.G. (2006). Development of line-detection algorithms for local positioning in densely seeded crops. *Canadian Biosystems Engineering*, 48(7), 19–29.
- Frizzelle, B.G., Moody, A. (2001). Mapping continuous distributions of land cover: A comparison of maximum-likelihood estimation and artificial neural networks, *Photogramm. Eng. Remote Sens.*, 67(6), 693–705.
- Fu, K.S., González, R.C., Lee, C.S.G. (1988). *Robótica: Control, detección, visión e inteligencia*. McGraw-Hill, Madrid.
- Fu, K.S., Mu, J.K. (1981). A survey on image segmentation. *Pattern Recognition* 13(1), 3-16.

- Gebhardt, S., Kaühbauch, W.A. (2007). A new algorithm for automatic *Rumex obtusifolius* detection in digital image using colour and texture features and the influence of image resolution. *Precision Agriculture*, 8(1), 1–13.
- Gebhardt, S., Schellberg, J., Lock, R., Kaühbauch, W.A. (2006). Identification of broad-leaved dock (*Rumex obtusifolius* L.) on grassland by means of digital image processing. *Precision Agriculture*, 7(3), 165–178.
- Gée, Ch., Bossu, J., Jones, G., Truchetet, F. (2008). Crop/weed discrimination in perspective agronomic images. *Computers and Electronics in Agriculture*, 60, 49-59.
- Gerhards, R., Oebel, H. (2006). Practical experiences with a system for site-specific weed control in arable crops using real-time image analysis and GPS-controlled patch spraying. *Weed Research*, vol. 46, no. 3, pp. 185-193.
- Giacinto, G., Roli, F., Bruzzone, L. (2000). Combination of neural and statistical algorithms for supervised classification of remote-sensing image, *Pattern Recognition Letters*, 21(5), 385-397.
- Gil, E. (2001). Situación actual y posibilidades de la agricultura de precisión. XXIX Conferencia Internacional de Mecanización Agraria (CIMA 2001)
- Golub, G., Van Loan, C. (1980). An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, 17, 883–893.
- Golub, G. (1973). Some modified matrix eigenvalue problems. *SIAM Review*, 15, 318–344.
- Gonzales-Barron, U., Butler, F. (2006). A comparison of seven thresholding techniques with the k-means clustering algorithm for measurement of bread-crumbs features by digital image analysis. *Journal of Food Engineering*, vol. 74, pages 268-278.
- Gonzalez, R.C., Woods, R.E. (2007). *Digital Image Processing*, 3rd Ed., Prentice-Hall, New Jersey, USA.
- Gottschalk, R., Burgos-Artizzu, X.P., Ribeiro, A., Pajares Miguelsanz, G., Sánchez, A. (2010). Real-time image processing for the guidance of a small agricultural field inspection vehicle. *International Journal of Intelligent Systems Technologies and Applications (IJISTA)*, vol.8, no.1-4, pp.434-443. January, 2010.
- GRDC (2012). Herbicide tolerance of PBA Gunyidi - a new Narrow Leafed Lupin variety. <http://www.grdc.com.au/Research-and-Development/GRDC-Update-Papers/2012/04/Herbicide-tolerance-of-PBA-Gunyidi-a-new-Narrow-Leafed-Lupin-variety> (accedido agosto 2014).
- Guerrero, J.M., Montalvo, M., Guijarro, M., Romeo, J., Pajares, G. (2014). Comparison of two new automatic techniques for exposure time control in agricultural images. Second RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture and forestry, mayo 21-23, Madrid (España) (<http://www.rhea-conference.eu/2014/>).

- Guerrero, J.M., Guijarro, M., Montalvo, M., Romeo, J., Emmi, L., Ribeiro, A. Pajares, G. (2013). Automatic expert system based on images for accuracy crop row detection in maize fields. *Expert Systems with Applications*, vol 40, issue 2, pages 656–664, doi: 10.1016/j.eswa.2012.07.073.
- Guerrero, J.M., Pajares, G., Montalvo, M., Romeo, J., Guijarro, M. (2012a). Support Vector Machines for crop/weeds identification in maize fields. *Expert Systems with Applications*, vol. 39, issue 12, pages 11149-11155, doi: 10.1016/j.eswa.2012.03.040.
- Guerrero, J.M., Montalvo, M., Romeo, J., Oliva, D., Guijarro, M., Pajares, G. (2012b). Detection of images with high CCD charge level due to indirect sunlight in maize fields. *International Conference of Agricultural Engineering*, julio 8-12, Valencia (España). ISBN-10: 84-615-9928-4 / ISBN-13: 978-84-615-9928-8 (<http://cigr.ageng2012.org>).
- Guerrero, J.M., Romeo, J., Emmi, L., Montalvo, M., Guijarro, M., Pajares, G., González-de-Santos, P. (2012c). Influence of the vision system pitch angle on crop and weeds detection accuracy. *First RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture*, septiembre 19-21, Pisa (Italia). ISBN-13: 978-88-6741-021-7 (<http://www.rhea-conference.eu/2012/>).
- Guerrero, J.M., Pajares, G., Guijarro, M. (2011). Técnicas de procesamiento de imágenes estereoscópicas. *Enlaces: revista del CES Felipe II*, ISSN 1695-8543, N°. 13.
- Guijarro, M., Guerrero, J.M., Montalvo, M., Romeo, J., Pajares, G. (2013). A new approach to solve image thresholding in precision agriculture. *9th European Conference on Precision Agriculture*. Lleida, Catalonia, Spain, July 7-11, poster. <http://www.ecpa2013.udl.cat/>
- Guijarro, M., Pajares, G., Riomoros, I., Herrera, P.J., Burgos-Artizzu, X.P., Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, 75, 75-83.
- Haapanen, R., Ek, A.R., Bauer, M.E., Findley, A.O. (2004). Delineation of forest/nonforest land use classes using nearest neighbor methods, *Remote Sensing Environment*, 89(3), 265-271.
- Hague, T., Tillett, N., Wheeler, H. (2006). Automated crop and weed monitoring in widely spaced cereals. *Precision Agriculture*, 1(1), 95–113.
- Hague, T., Marchant, J.A., Tillett, D. (1997). A system for plant scale husbandry. *Precision Agriculture*, 635–642.
- Hanmandlu, M., Madasu, V.K., Vasikarla, S. (2004). A Fuzzy Approach to Texture Segmentation, in: *Proc. of the IEEE International Conference on Information Technology: Coding and Computing (ITCC'04)*, The Orleans, Las Vegas, Nevada, USA, 636-642.
- Hansen, M., Dubayah, R., Defries, R. (1996). Classification trees: an alternative to traditional land cover classifiers, *Int. J. Remote Sensing*, 17(5), 1075-1081.

- Haralick, R.M., Shanmugam, K., Dinstein, I. (1973). Texture features for image classification, *IEEE Trans. Systems Man Cybernet.* SMC-3, 610-621.
- Haykin, S. (1994). *Neural Networks: A comprehensive Foundation*, McMillan College Publishing Co., New York.
- Helmers, H., Schellenberg, M. (2003). CMOS vs. CCD sensors in speckle interferometry, *Optics & Laser Technology*, 35, 587-595.
- Hough, P.V.C. (1962). A method and means for recognizing complex patterns. U.S. Patent Office No. 3069654.
- Huang, C.L. (1999). Pattern image segmentation using modified Hopfield model, *Pattern Recognition Letters*, 13, 345-353.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(2-Jan), 195-213.
- Huete, A.R., Liu, H.Q., Batchily, K., vanLeeuwen, W. (1997). A comparison of vegetation indices global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59(3), 440-451.
- Huete, A.R., Justice, C., Liu, H.Q. (1994). Development of vegetation and soil indices for MODIS-EOS. *Remote Sensing of Environment*, vol 49, issue 3, pages 224-234, doi: [http://dx.doi.org/10.1016/0034-4257\(94\)90018-3](http://dx.doi.org/10.1016/0034-4257(94)90018-3).
- Huete, A.R. (1988). A Soil-adjusted vegetation index. *Remote Sensing of Environment*, 25:295-309.
- ISCAR (2014). *Ingeniería de Sistemas, Control, Automatización y Robótica*, aprobado por la Universidad Complutense de Madrid con número de referencia 910793. <http://www.ucm.es/grupos/gi477> (accedido agosto 2014).
- Jacobson, R.E., Ray, S.F., Attridge, G.G., Axford, N.R. (2000). *The manual of photography: Photographic and digital imaging*. Ninth Edition. Focal Press Oxford.
- Ji, R., Qi, L. (2011). Crop-row detection algorithm based on random Hough transformation. *Mathematical and Computer Modelling*, 54, 1016–1020.
- Jiji, G.W., Ganesan, L. (2008). A new approach for unsupervised segmentation, *Applied Soft Computing Journal*, doi:10.1016/j.asoc.2008.06.002.
- Johannsen, G., Bille, J. (1982). A threshold selection method using information measures, in *Proceedings, 6th Int. Conf. Pattern Recognition*, Munich, Germany, pp. 140-143.
- Jones, G., Gée, Ch., Truchetet, F. (2009a). Modelling agronomic images for weed detection and comparison of crop/weed discrimination algorithm performance. *Precision Agriculture*, 10, 1–15.

- Jones, G., Gée, Ch., Truchetet, F. (2009b). Assessment of an inter-row weed infestation rate on simulated agronomic images. *Computers and Electronics in Agriculture*, 67, 43–50.
- Kapur, J.N., Sahoo, P.K., Wong A.K.C. (1985). A new method for gray-level picture thresholding using the entropy of the histogram, *Comput. Vision Graphics Image Process.* 29, 273-285.
- Kataoka, T., Kaneko, T., Okamoto, H., Hata, S. (2003). Crop growth estimation system using machine vision. In *The 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*.
- Katz, Y.H. (1965). Pattern recognition of meteorological satellite cloud photography, *Proceedings, Third Symp. on Remote Sensing of Environment*, pp 173-214.
- Kendall, M. (1955). *Rank correlation methods* (2nd ed.). London: Charles Griffin and Company.
- Kendall, M. (1938). A New Measure of Rank Correlation. *Biometrika* 30 (1–2): 81–89. doi:10.1093/biomet/30.1-2.81. JSTOR 2332226.
- Kirby, R.L., Rosenfeld, A. (1979). A note on the use of (gray level, local average gray level) space as an aid in thresholding selection, *IEEE Trans. Systems Man Cybernet.* SMC-9, 860-864.
- Kirk, K., Andersen, H.J., Thomsen, A.G., Jørgensen, J.R. (2009). Estimation of leaf area index in cereal crops using red-green images. *Biosystems Engineering*, 104, 308–317.
- Kise, M., Zhang, Q. (2008). Development of a stereovision sensing system for 3D crop row structure mapping and tractor guidance. *Biosystems Engineering*, 101, 191–198.
- Kise, M., Zhang, Q., Rovira-Más, F. (2005). A stereovision-based crop row detection method for tractor-automated guidance. *Biosystems Engineering*, 90 (4), 357–367.
- Kittler, J., Illingworth, J. (1986). Minimum error thresholding. *Pattern Recognition*, 19(1), 41–47.
- Kohler, R. (1981). A segmentation system based on thresholding, *Comput. Graphics Image Process.* 15, 319-338.
- Komala, J., Punitha M. (2009a). Computation of Merging Points in Skeleton Based Images. *Global Journal Of computer science and technology GJCST*. Volume 9 Number 4 September 2009 Issue. Print ISSN: 0975 - 4350, Online ISSN : 0975 - 4172.
- Komala, J., Punitha M. (2009b). A Survey on skeletons in digital image processing. *International conference proceedings of IEEE Computer Society*: 260-269.
- Komala, J., Punitha M. (2008). A Survey on skeletonization in digital image processing. *International conference proceedings of Managing Next Generations Software Applications 08*, Sponsored by CSIR New Delhi: 825-839.

- Kruegle, H. (2006). *CCTV Surveillance (Second Edition)*, Butterworth-Heinemann, Burlington, ISBN 978-0-7506-7768-4, Chapter 4 - Lenses and Optics, Pages 71-107, <http://dx.doi.org/10.1016/B978-075067768-4/50007-1>.
- Kruskal, W.H. (1958). Ordinal Measures of Association. *Journal of the American Statistical Association* 53 (284): 814–861. doi:10.2307/2281954. JSTOR 2281954. MR 100941.
- Kumar, S., Ghosh, J., Crawford, M.M. (2001). Best-bases feature extraction for pairwise classification of hyperspectral data, *IEEE Trans. Geoscience and Remote Sensing*, 39(7), 1368-1379.
- Kumar, A.S., Majumder, K.L. (2001). Information fusion in tree classifiers, *Int. J. Remote Sensing*, 22(5), 861-869.
- Lee, J.Y., Wang, Y.H., Lai, L.J., Lin, Y.J., Chang, Y.H. (2011). Development of an auto-focus system based on the moiré method. *Measurement*, 44, 1793-1800.
- Leemans, V., Destain, M.F. (2006). Application of the Hough transform for seed row location using machine vision. *Biosystems Engineering*, 94(3), 325–336.
- Lillo, A., Motta, G., Storer, J. A. (2007). Supervised segmentation based on texture signatures extracted in the Frequency Domain, in: J. Martí, J.M. Benedí. A.M. Mendonça, J. Serrat (Eds.), *Pattern Recognition and Image Analysis*, Springer Verlag, Berlin, LNCS 4477, Part I, 89-96.
- Lim, Y.W., Lee, S.U. (1990). On the color image segmentation algorithm based on the thresholding and the fuzzy c-means technique, *Pattern Recognition*, 23(9), 935-952.
- Ling, P.P., Ruzhitsky, V.N. (1996). Machine vision techniques for measuring the canopy of tomato seedling. *Journal Agricultural Engineering Research*, 65(2), 85–95.
- Litwiller, D. (2001). CCD Vs. CMOS: Facts and Fiction. *Photonics Spectra*. Online: https://www.teledynedalsa.com/public/corp/Photonics_Spectra_CCDvsCMOS_Litwiller.pdf (accedido agosto 2014).
- Lo, Y.S., Pei, S.C. (1999). Color image segmentation using local histogram and selforganizing of Kohonen feature map, in: *Proc. Int. Conf. on Image Processing*, Kobe, Japan, 232-239.
- López-Granados, F. (2011). Weed detection for site-specific weed management: mapping and real-time approaches. *Weed Research*, 51, 1–11.
- López-Orozco, J.A. (1999). *Integración y Fusión multisensorial*. Tesis Doctoral, Universidad Complutense, Madrid.
- Low, A. (1991). *Introductory Computer Vision and Image Processing*. McGraw-Hill, New York.

- Luscier, J.D., Thompson, W.L., Wilson, J.M., Gorham, B.E., Dragut, L.D. (2006). Using digital photographs and object-based image analysis to estimate percent ground cover in vegetation plots. *Frontiers in Ecology and the Environment*, 4(8), 408–413.
- Magnan, P. (2003). Detection of visible photons in CCD and CMOS: A comparative view. *Nuclear Instruments and Methods in Physics Research A* 504, 199-212.
- Marchant, J. (1996). Tracking of row structure in three crops using image analysis. *Computers and Electronics in Agriculture*, 15, 161–179.
- Marcos, A.G., Martínez de Pisón Ascacíbar, F.J., Pernía Espinoza A.V., Alba Elías, F., Castejón Limas, M., Ordieres Meré, J., Vergara, E., (2006). *Técnicas y Algoritmos Básicos de Visión Artificial*, First Edition ed. Vol. 1. Universidad de la Rioja. 91 pp.
- Mas, J.F., Flores, J.J. (2008). The application of artificial neural networks to the analysis of remotely sensed data, *Int. J. Remote Sensing*, 29(3), 617-663.
- Mason, D., Lauder, I. J., Rutoritz, D., Spowart, G. (1975). Measurement of C-Bands in human chromosomes, *Comput. Biol. Med.* 5, 179-201.
- Melgani, F., Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines, *IEEE Trans. Geosci. Remote Sensing*, 42(8), 1178–1790.
- Meyer, G.E., Camargo-Neto, J. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture*, 63, 282–293.
- Meyer, G.E., Camargo-Neto, J., Jones, D.D., Hindman, T.W. (2004). Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Computers and Electronics in Agriculture*, 42, 161–180.
- Meyer, G.E., Hindman, T.W., Lakshmi, K. (1998). *Machine vision detection parameters for plant species identification*. Bellingham, WA: SPIE.
- MicroStrain (2014). Inertial Sensors 3DM-GX3-35 con GPS. Disponible online: <http://www.microstrain.com/inertial/3dm-gx3-35> (accedido agosto 2014).
- Mohaddamzadeh A., Bourbakis, N. (1997). A fuzzy region growing approach for segmentation of color images, *Pattern Recognition*, 30(6), 867-881.
- Montalvo, M., Guerrero, J.M., Guijarro, M., Romeo, J., Pajares, G. (2014). Exposure time control in agricultural images using a color adaptive method. Second RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture and forestry, mayo 21-23, Madrid (España) (<http://www.rhea-conference.eu/2014/>).
- Montalvo, M., Guerrero, J.M., Romeo, J., Emmi, L., Guijarro, M., Pajares, G. (2013a). Automatic expert system for weeds/crops identification in images from maize fields. *Expert Systems with Applications*, vol 40, issue 1, pages 75–82, doi: 10.1016/j.eswa.2012.07.034.

- Montalvo, M., Guerrero, J.M., Romeo, J., Guijarro, M., de la Cruz, J.M., Pajares, G. (2013*b*). Acquisition of Agronomic Images with Sufficient Quality by Automatic Exposure Time Control and Histogram Matching. Lecture Notes in Computer Science. Advanced Concepts for Intelligent Vision Systems (ACIVS'13) (Jacques Blanc-Talon, Andrzej Kasinski, Wilfried Philips, Dan Popescu, Paul Scheunders, Eds.), 8192, 37-48. Springer International Publishing Switzerland. October 28-31, Poznan, Poland.
- Montalvo, M., Pajares, G., Guerrero, J.M., Romeo, J., Guijarro, M., Ribeiro, A., Ruz, J.J., de la Cruz, J.M. (2012*a*). Automatic detection of crop rows in maize fields with high weeds pressure. *Expert Systems with Applications*, 39(15), 11889–11897, doi: 10.1016/j.eswa.2012.02.117.
- Montalvo, M., Guerrero, J.M., Romeo, J., Oliva, D., Guijarro, M., Pajares, G. (2012*b*). Unsupervised learning for crop/weeds discrimination in maize fields with high weeds densities. *International Conference of Agricultural engineering*, julio 8-12, Valencia (España). ISBN-10: 84-615-9928-4 / ISBN-13: 978-84-615-9928-8 (<http://cigr.ageng2012.org>).
- Montalvo, M., Guerrero, J.M., Guijarro, M., Romeo, J., Herrera, P.J., Ribeiro, A., Pajares, G. (2011). Camera System geometry for site specific treatment in precision agriculture. In *Proc. First RHEA Workshop, Robotics and associated High-technologies and Equipment for Agriculture* (G. Rabatel and P. González de Santos), September 9, Montpellier (France).
- Monteith, J.L. (1981). Evaporation and surface temperature. *Quart. J. Roy. Meteorolog. Soc.*, 107, 1–27.
- Morrin, T.H. (1974). A black-white representation of a gray-scale picture, *IEEE Trans. Comput.* 23, 184-186.
- Muñoz-Marí, J., Bruzzone, L., Camps-Valls, G. (2007). A Support Vector Domain Description Approach to Supervised Classification of Remote Sensing Images, *IEEE Trans. Geosci. Remote Sens.*, 45(8), 2683–2692.
- Nakagawa, Y., Rosenfeld, A. (1979). Some experiments on variable thresholding, *Pattern Recognition* 11, 191-204.
- National Instruments (2014). Disponible online: <http://spain.ni.com/> (accedido agosto 2014).
- Neto, J.C. (2004). A combined statistical-soft computing approach for classification and mapping weed species in minimum tillage systems. Lincoln, NE: University of Nebraska.
- Newsam, S., Wang, L. Bhagavathy, S., Manjunath, B.S. (2004). Using Texture to analyze and Manage Large Collections of Remote Sensed image and Video Data, *Applied Optics*, 43(2), 210-217.
- Nixon, M., Aguado, A. (2002). Feature extraction & image processing. Oxford: Newnes.
- Olsen, H.J. (1995). Determination of row position in small-grain crops by analysis of video images. *Computers and Electronics in Agriculture*, 12, 147–162.

- Otsu, N. (1979). A threshold selection method from gray-level histogram. *IEEE Transactions on Systems Man and Cybernetics*, 9, 62–66.
- Onyango, C.M., Marchant, J.A. (2003). Segmentation of row crop plants from weeds using colour and morphology. *Computers and Electronics in Agriculture*, 39, 141–155.
- Pajares, G., Guerrero, J.M., Romeo, J., Sánchez-Benítez, D., Montalvo, M., Herrera, P.J., Guijarro, M., de la Cruz, J.M., Ruz, J.J. (2012). Deterministic Simulated Annealing applied to Computer Vision tasks. *Journal of High Performance Computing*. (<http://bioinfopublication.org/viewhtml.php?artid=BIA0000418>) *Bioinfo Publications*, Volume 2, Issue 1, pp.-20-23, ISSN: 2230–7192 & E-ISSN: 2230–7206.
- Pajares, G., Cruz, J.M. (2007). *Visión por Computador: Imágenes digitales y aplicaciones*, 2ª ed., RA-MA, Madrid.
- Pajares, G., Cruz, J.M. (2002). Clasificación de Texturas Naturales mediante K-Means, *Revista Electrónica de Visión por Computador* (ISSN 1575-5258), no. 6, 1-18, Ed. Centre de Visió per Computador, Universitat Autònoma, Barcelona.
- Pajares, G., Cruz, J.M., Moreno, V. (2002). Clasificación de texturas naturales mediante agrupamiento borroso, *Ingeniería Civil*, Centro de Estudios y Experimentación de Obras Públicas (CEDEX), Ministerio de Fomento, 127, 83-89.
- Pajares, G., Cruz, J.M., Moreno, V. (2001). Clasificación de texturas mediante redes neuronales, *Ingeniería Civil*, Centro de Estudios y Experimentación de Obras Públicas (CEDEX), Ministerio de Fomento, 123, 61-69.
- Panda, D.P. (1977). *Segmentation of FLIR Images by Pixel Classification*, University of Maryland Computer Science Center.
- Paola, J.D., Schowengerdt, R.A. (1995). A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery, *Int. J. Remote Sensing*, 16(16), 3033-3058.
- Paruelo, J.M., Epstein, H.E., Lauenroth, W.K., Burke, I.C. (1997). ANPP estimates from NDVI for the central grasslands region of the U.S. *Ecology* 78: 953-958.
- Pavlidis, T. (1977). *Structural Pattern Recognition*, Springer-Verlag, New York.
- Peleg, S. (1980). A new probabilistic relaxation scheme, *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-2*, issue 4, 362-369.
- Peña J.M., Torres-Sánchez J., de Castro A.I., Kelly M., López-Granados, F. (2013) Weed Mapping in Early-Season Maize Fields Using Object-Based Analysis of Unmanned Aerial Vehicle (UAV) Images. *PLoS ONE* 8(10): e77151. doi:10.1371/journal.pone.0077151
- Peña, J.M., Kelly, M., de Castro, A.I., Lopez-Granados, F. (2012). Discrimination of crop rows using object-based approaches in UAV images for early site-specific weed management in maize fields. *Proc. of the 1st International Conference on Robotics and Associated High Technologies and Equipment for Agriculture*, Pisa, Italy, 19–21 September 2012, 249–254.

- Pérez, A.J., López, F., Belloch, J.V., Christensen, S. (2000). Color and shape analysis techniques for weed detection in cereal fields. *Computer and Electronics in agriculture*, vol. 25, pages 197-212.
- Pitas, I. (2000). *Digital Image Processing Algorithms and Applications*. John Wiley & Sons. ISBN-10: 0471377392 / ISBN-13: 978-0471377399
- Pla, F., Sanchiz, J.M., Marchant, J.A., Brivot, R. (1997). Building perspective models to guide a row crop navigation vehicle. *Image and Vision Computing*, 15, 465–473.
- Ponti, M.P. (2013). Segmentation of Low-Cost Remote Sensing Images Combining Vegetation Indices and Mean Shift. *Geoscience and Remote Sensing Letters, IEEE* (Volume:10 , Issue: 1).
- Prewitt, J.M.S., Mendelsohn, M.L. (1966). The analysis of cell images, in *Ann. New York Academy Science Vol 128*, pp 1035-1053, New York.
- Puig, D., García, M.A. (2006). Automatic texture feature selection for image pixel classification, *Pattern Recognition*, 39(11), 1996-2009.
- Pun, T. (1981). Entropic thresholding: A new approach, *Comput. Vision Graphics Image Process.* 16, 210-239.
- Pun, T. (1980). A new method for gray-level picture thresholding using the entropy of the histogram, *Signal Processing* 2(3), 223-237.
- Pyun, K., Won, C.S., Lim, J., Gray, R.M. (2002). Texture classification based on multiple Gauss mixture vector quantizer, in: *Proc. of ICME*, 501-504.
- Rabatel, G., Gorretta, N., Labbé, S. (2014). Getting simultaneous red and near-infrared band data from a single digital camera for plant monitoring applications: Theoretical and practical study. *Biosystems Engineering*, 117, 2-14.
- Randen, T., Husøy, J.H. (1999). Filtering for Texture Classification: A Comparative Study, *IEEE Trans. Pattern Analysis Machine Intell.*, 21(4), 291-310.
- Rashidian, B., Fox, E. (2011). The evolution of CMOS technology (Teledyne Dalsa). https://www.teledynedalsa.com/public/mv/appnotes/EvolutionofCMOS_Technology_wp.pdf (accedido agosto 2014).
- Reid, J.F., Searcy, S.W. (1987). Vision-based guidance of an agricultural tractor. *IEEE Control Systems*, 7(12), 39–43.
- RHEA (2014). Robot Fleets for Highly Effective Agriculture and Forestry Management. NMP-CP-IP 245986-2 RHEA. Disponible online: <http://www.rhea-project.eu/> (accedido agosto 2014).
- Ribeiro, A., Fernández-Quintanilla, C., Barroso, J., García-Alegre, M.C. (2005). Development of an image analysis system for estimation of weed, *Proceedings of the 5th European Conference on Precision Agriculture (5ECPA)* 169–174.
- Richardson, A.J., Wiegand, C.L. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, vol. 43, pages 1541-1552.

- Ridler, T. W., Calvard, S. (1978). Picture thresholding using an iterative selection method. *IEEE Transactions on Systems, Man and Cybernetics*, 8(8), 630–632.
- Romeo, J., Benítez, A., Guerrero, J.M., Montalvo, M., Pajares, G. (2014). Tractor guidance in maize fields using computer vision and geometry. Second RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture and forestry, mayo 21-23, Madrid (España) (<http://www.rhea-conference.eu/2014/>).
- Romeo, J., Pajares, G., Montalvo, M., Guerrero, J.M., Guijarro, M., de la Cruz, J.M. (2013a). A new Expert System for greenness identification in agricultural images. *Expert Systems with Applications*, Volume 40, Issue 6, Pages 2275-2286 doi: 10.1016/j.eswa.2012.10.033.
- Romeo, J., Guerrero, J.M., Montalvo, M., Emmi, L., Guijarro, M., Gonzalez-de-Santos, P., Pajares, G. (2013b). Camera sensor arrangement for crop/weeds detection accuracy in agronomic images. *Sensors*, 13, 4348-4366, doi: 10.3390/s130404348.
- Romeo, J., Pajares, G., Montalvo, M., Guerrero, J.M., Guijarro, M., Ribeiro, A. (2012a). Crop row detection in maize fields inspired on the human visual perception. *The Scientific World Journal*, 2012. <http://dx.doi.org/10.1100/2012/484390>. Article ID 484390, 10 pages.
- Romeo, J., Oliva, D., Guerrero, J.M., Montalvo, M., Guijarro, M., Pajares, G. (2012b). Plants discrimination in maize fields applying human perception to digital images. *International Conference of Agricultural Engineering*, julio 8-12, Valencia (España). ISBN-10: 84-615-9928-4 / ISBN-13: 978-84-615-9928-8 (<http://cigr.ageng2012.org>).
- Rouse, J.W., Haas, R.H., Schell J.A., Deering, D.W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. *Proc. Third ERTS-1 Symposium*, NASA Goddard, NASA SP-351 pp. 309-317.
- Rovira-Mas, F., Zhang, Q., Hansen, A.C. (2010). *Mechatronics and Intelligent Systems for Off-road Vehicles*, Springer-Verlag. ISBN: 978-1-84996-467-8 (Print) 978-1-84996-468-5 (Online).
- Rovira-Más, F., Zhang, Q., Reid, J.F. (2008). Stereo vision three-dimensional terrain maps for precision agriculture. *Computers and Electronics in Agriculture*, 60(2), 133–143.
- Rovira-Más, F., Zhang, Q., Reid, J.F., Will, J.D. (2005). Hough-transform-based vision algorithm for crop row detection of an automated agricultural vehicle. *Journal of Automobile Engineering Part D*, 219, 999–1010.
- Rovira-Mas F., Zhang Q., Reid J.F. and Will J.D. (2003). Machine vision based automated tractor guidance. *Int. J. Smart Eng. Syst. Des.*, 5, 467–480
- Rosenfeld, A., De La Terre, P. (1983). Histogram concavity analysis as an aid in threshold selection, *IEEE Trans. Systems Man Cybernet.* SMC-13, 231-235.
- Rosenfeld, A., Smith, R.C. (1981). Thresholding using relaxation, *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-3, 598-606.

- Rud, R., Shoshany, M., Alchanatis, V., Cohen, Y. (2006). Application of spectral features' ratios for improving classification in partially calibrated hyperspectral imagery: a case study of separating Mediterranean vegetation species, *Journal Real-Time Image Processing*, 1, 143-152.
- Ruiz-Ruiz, G., Gómez-Gil, J., Navas-Gracia, L.M. (2009). Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm (EASA). *Computers and Electronics in Agriculture*, 68, 88–96.
- Rumpf, T., Römer, C., Weis, M., Sökefeld, M., Gerhards, R., Plümer, L. (2012). Sequential support vector machine classification for small-grain weed species discrimination with special regard to *Cirsium arvense* and *Galium aparine*. *Computers and Electronics in Agriculture*, vol. 80, pp. 89-96.
- Sahoo, P.K., Soltani, S., Wong A.K.C., Chen Y.C. (1988). A survey of thresholding techniques. *Computer Vision, Graphics and Image Processing*, 41, 233-260.
- Sainz-Costa, N., Ribeiro, A., Burgos-Artizzu, X.P., Guijarro, M., Pajares, G. (2011). Mapping wide row crops with video sequences acquired from a tractor moving at treatment speed. *Sensors*, 11, 7095–7109.
- Schneider Kreuznach (2014a). C-Mount Lenses compact series 1". Disponible on-line: <http://www.schneiderkreuznach.com/en/industrial-solutions/lenses-and-accessories/products/1-oe16mm-lenses/compact-lenses/cinegon-1910/> (accedido agosto 2014)
- Schneider Kreuznach (2014b). Tipps and Tricks. Disponible online: <http://www.schneiderkreuznach.com/en/photo-imaging/product-field/b-w-fotofilter/products/filtertypes/special-filters/486-uvir-cut/> (accedido agosto 2014).
- Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall's Tau. *Journal of American Statistical Association*, 63, 1379–1389.
- Seong, J.C., Ustry, E.L. (2001). Fuzzy image classification for continental scale multitemporal NDVI series images using invariant pixels and an image stratification method, *Photogrammetric Engineering and Remote Sensing*, 67(3), 287-294.
- Serra, J. (1982). *Image analysis and mathematical morphology*, Academic Press, New-York, ISBN 0-12-637240-3.
- Shapiro, L.G., Stockman, G.C. (2001) *Computer Vision*. Prentice Hall. First edition. ISBN 0-13-030796-3.
- Shrestha, D.S., Steward, B.L., Birrell, S.J. (2004). Video processing for early stage maize plant detection. *Biosystem Engineering*, 89(2), 119–129.
- Shi, Y.Q., Sun, H. (2000). *Image and Video Compression for Multimedia Engineering*, CRC Press, ISBN 0-8493-3491-8.
- Singh, S., Haddon, J., Markou, M. (2001). Nearest-neighbour classifiers in natural scene analysis, *Pattern Recognition*, 34, 1601-1612.

- SINTEF (2014). Disponible online: <http://www.sintef.no/home/Press-Room/Research-News/Focused-warfare-on-weeds/> (accedido agosto 2014).
- Slaughter, D.C., Giles, D.K., Downey, D. (2008). Autonomous robotic weed control systems: A review. *Computers and Electronics in Agriculture*, 61, 63–78.
- Smith, G., Burns, I. (1997). Measuring texture classification algorithms, *Pattern Recognition Letters*, 18, 1495-1501.
- Sneath, P., Sokal, R. (1973). *Numerical Taxonomy: The Principle and Practice of Numerical Classification*. W.H. Freeman, San Francisco
- Søgaard, H.T., Olsen, H.J. (2003). Determination of crop rows by image analysis without segmentation. *Computers and Electronics in Agriculture*, 38, 141–158.
- Southwell, R. (1946). *Relaxation Methods in Theoretical Physics*, Oxford Univ. Press (Clarendon), London.
- Southwell, R. (1940). *Relaxation Methods in Engineering Science, A Treatise on Approximate Computation*, Oxford Univ. Press, London.
- Stoll, A. (2003). Automatic operation planning for GPS-guided machinery, in *Proceedings of the Fourth European Conference on Precision Agriculture*, Berlin, edited by J.V. Stafford & A. Werner, pp. 657-664.
- SVS-VISTEK (2014). The Focal Point of Machine Vision. Disponible online: <http://www.svs-vistek.com/intl/en/index.php> (accedido agosto 2014).
- Tang, L., Tian, L.F., Steward, B.L. (2000). Color image segmentation with genetic algorithm for in-field weed sensing. *Transactions of the ASAE*. 43(4): 1019-1027.
- Tellaache, A., Pajares, G., Burgos-Artizzu, X.P., Ribeiro, A. (2011). A computer vision approach for weeds identification through support vector machines. *Applied Soft Computing*, (ISSN 1568-4946) vol. 11(1), pp. 908–915.
- Tellaache, A., Burgos-Artizzu, X.P., Pajares, G., Ribeiro, A. (2008a). A vision-based method for weeds identification through the Bayesian decision theory. *Pattern Recognition*, 41, 521–530.
- Tellaache, A., Burgos-Artizzu, X.P., Pajares, G., Ribeiro, A., Fernández-Quintanilla, C. (2008b). A new vision-based approach to differential spraying in precision agriculture. *Computers and Electronics in Agriculture*, 60(2), 144–155.
- Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. In I, II, and III, *Nederl. Akad. Wetensch. Proc.*, 53, 386–392, 521–525 and 1397–1412.
- TheMathworks (2014). Disponible online: <http://www.mathworks.com/> (accedido agosto 2014)
- Tian, L.F., Slaughter, D.C. (1998). Environmentally adaptive segmentation algorithm for outdoor image segmentation. *Computers and Electronics in Agriculture*, 21, 153-168.

- Tottrup, C. (2007). Forest and Land Cover Mapping in a Tropical Highland Region, *Photogrammetric Engineering and Remote Sensing*, 73(9), 1057-1066.
- Trauwaert, E., Kaufman, L., Rousseeuw, P. (1991). Fuzzy Clustering Algorithms Based on the Maximum Likelihood Principle, *Fuzzy Sets and Systems*, 42, 213-227.
- Tsai, W. (1985). Moment-preserving thresholding: A new approach, *Comput. Vision Graphics Image Process.* 29, 377-393.
- Tso, B.C.K., Mather, P.M. (1999). Classification of multisource remote sensing imagery using a genetic algorithm and Markov random fields, *IEEE Trans. Remote Sensing*, 37(3), 1255-1260.
- Tucker, C.J., Vanpraet, C.L., Sharman M.J., Van Ittersum, G. (1985). Satellite remote sensing of total herbaceous biomass production in the senegalese sahel: 1980–1984. *Remote Sensing of Environment*, vol. 17, issue 3, 233-249, doi: [http://dx.doi.org/10.1016/0034-4257\(85\)90097-5](http://dx.doi.org/10.1016/0034-4257(85)90097-5).
- Ullman, J.R. (1974). Binarization using associative addressing, *Pattern Recognition* 6(2), 127-135.
- Vapnik, V.N. (2000). *The nature of statistical learning theory*. New York: Springer-Verlag.
- Vioix, J.B., Douzals, J.P., Truchetet, F., Assemat, L., Guillemin, J.P. (2002). Spatial and spectral method for weeds detection and localization. *EURASIP JASP*, 7, 679–685.
- Virginia, R.A., Wall, D.H. (2001). Ecosystem Function, Principles of. In *Encyclopedia of Biodiversity*, (ed. Simon A. Levin), pp. 345-352. New York: Elsevier.
- Wagner, T. (1999). Texture Analysis, in: B. Jahne, H. Haussecker, P. Geisser (Eds.), *Handbook of Computer Vision and Applications*, Academic Press, San Diego, cap. 2 (Signal Processing and Pattern Recognition).
- Wang, S., Haralick, R.M. (1984). Automatic multithreshold selection, *Comput. Vision Graphics Image Process.* 25, 46-67.
- Watanabe, S., the CYBEST group (1974). An automated apparatus for cancer processing CYBEST, *Comput. Vision Graphics Image Process.* 3, 350-358.
- Webots (2014). Cyberbotics. Webots: Robot Simulator. Disponible online <http://www.cyberbotics.com/> (accedido agosto 2014).
- Weszka, J.S., Rosenfeld, A. (1979). Histogram modification for threshold selection, *IEEE Trans. Systems Man Cybernet.* SMC-9, 38-51.
- Weszka, J.S. (1978). A survey of threshold selection techniques. *Computer Graphics Image Processing*, vol. 7, 259-265.
- Weszka, J.S., Rosenfeld, A. (1974). *Threshold Selection 4*, University of Maryland Computer Science Center.
- Weszka, J.S., Nagel, R.N., Rosenfeld, A. (1974). A threshold selection technique, *IEEE Trans. Comput.* C-23, 1322-1326.

- Weszka, J.S., Veretun, J.A., Rosenfeld, A. (1973). A Technique for Facilitating Threshold Selection for Objects Extraction from Digital Pictures, University of Maryland Computer Science Center.
- Woebbecke, D.M., Meyer, G.E., von Bargen, K., Mortensen, D.A. (1995). Shape features for identifying young weeds using image analysis. *Transactions on American Society of Agricultural Engineering* 38 (1), 271–281.
- Wolfe, R.N. (1969). A dynamic thresholding scheme for quantization of scanned image, in *Proceedings, Automatic Pattern Recognition*, pp. 143-162.
- Wu A.Y., Rosenfeld, A. (1982). Threshold selection using quadtree, *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-4*, 90-94.
- Xie, J., Tsui, H.T. (2004). Image Segmentation based on maximum-likelihood estimation and optimum entropy-distribution (MLE-OED), *Pattern Recognition Letters*, 25(10), 1133-1141.
- Yu, H., Li, M., Zhang, H.J., Feng, J. (2002). Color texture moments for content-based image retrieval, in: *Proc. International Conference on Image Processing*, 3, 24-28.
- Zaman, Q., Esau, T., Schumann, A., Percival, D.C., Chang, Y.K., Read, S.M., Farooque, A.A. (2011). Development of prototype automated variable rate sprayer for real-time spot-application of agrochemicals in wild blueberry fields. *Computers and Electronics in Agriculture*, 76(2), 175–182.
- Zhang, Z., Kodagoda, S., Ruiz, D., Katupitiya, J., G. Dissanayake, G. (2010). Classification of Bidens in wheat farms. *International Journal of Computer Applications in Technology (IJCAT)*, Vol.39 , No. 1/2/3, August, pp 123-129.
- Zhang, J., Foody, G.M. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery, *Int. J. Remote Sensing*, 19(14), 2721-2738.
- Zheng, L., Shi, D., Zhang, J. (2010). Segmentation of green vegetation of crop canopy images based on mean shift and Fisher linear discriminate. *Pattern Recognition Letters*, 31(9), 920–925.
- Zheng, L., Zhang, J., Wang, Q. (2009). Mean-shift-based color segmentation of images containing green vegetation. *Computers and Electronics in Agriculture*, 65, 93–98.
- Zucker, S., Hummel, R., Rosenfeld, A. (1977). An application of relaxation labelling to line and curve enhancement, *IEEE Trans. Comput. C-26*, 394-403.